```
                    GenCore version 4.5
            Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:         March 23, 2000, 19:36:16 ; Search time 3853.81 Seconds
                                            (without alignments)
                                            199.863 Million cell updates/sec

Title:          US-09-183-789-1
Perfect score:  204
Sequence:       1 gatcgttgtaccttcaag..........gacactgtagtttggtgc 204

Scoring table:  IDENTITY_NUC
Searched:       4538634 seqs, 1887831982 residues
Database :      EST:*
Word size :     0

Number of hits that pass the threshold  :  9077268
```

```
 1: em_est1:*        50: gb_est31:*
 2: em_est2:*        51: gb_est32:*
 3: em_est3:*        52: em_est20:*
 4: em_est4:*        53: em_est21:*
 5: em_est5:*        54: em_est22:*
 6: em_est6:*        55: em_est23:*
 7: em_est7:*        56: em_est24:*
 8: em_est8:*        57: em_est25:*
 9: em_est9:*        58: em_est26:*
10: em_est10:*       59: gb_est33:*
11: em_est11:*       60: gb_est34:*
12: em_est12:*       61: gb_est35:*
13: em_est13:*       62: gb_est36:*
14: em_est14:*       63: gb_est37:*
15: em_est15:*       64: gb_est38:*
16: em_est16:*       65: em_est27:*
17: em_est17:*       66: em_est28:*
18: em_est18:*       67: em_est29:*
19: em_est19:*       68: em_est30:*
20: gb_est1:*        69: gb_est39:*
21: gb_est2:*        70: gb_est40:*
22: gb_est3:*        71: gb_est41:*
23: gb_est4:*        72: gb_est42:*
24: gb_est5:*        73: gb_est43:*
25: gb_est6:*        74: gb_est44:*
26: gb_est7:*        75: em_est31:*
27: gb_est8:*        76: em_est32:*
28: gb_est9:*        77: em_est33:*
29: gb_est10:*       78: em_est34:*
30: gb_est11:*       79: gb_gss1:*
31: gb_est12:*       80: gb_gss2:*
32: gb_est13:*       81: gb_gss3:*
33: gb_est14:*       82: gb_gss4:*
34: gb_est15:*       83: gb_gss5:*
35: gb_est16:*       84: em_gss1:*
36: gb_est17:*       85: em_gss2:*
37: gb_est18:*       86: em_gss3:*
38: gb_est19:*       87: gb_gss6:*
39: gb_est20:*       88: gb_gss7:*
40: gb_est21:*       89: gb_gss8:*
41: gb_est22:*       90: gb_gss9:*
42: gb_est23:*       91: gb_gss10:*
43: gb_est24:*       92: em_gss2:*
44: gb_est25:*       93: em_gss3:*
45: gb_est26:*       94: em_gss5:*
46: gb_est27:*       95: em_gss6:*
47: gb_est28:*       96: em_gss8:*
48: gb_est29:*       97: em_gss9:*
49: gb_est30:*       98: em_gss10:*
                     99: em_gss11:*
                    100: gb_gss10:*
                    101: em_gss12:*
                    102: gb_gss12:*
                    103: gb_gss13:*
                    104: gb_gss14:*
                    105: gb_gss15:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                              SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	184.2	90.3	515	74	AW197564	AW197564 xm44d05.x
	2	179.2	87.8	510	63	AW004014	AW004014 wq94h07.x
c	3	174.2	85.4	505	49	AI624830	AI624830 ts71f12.x
c	4	110	53.9	330	29	AA155054	AA155054 mr91b11.r
c	5	108.4	53.1	425	42	AI654710	AI654710 wb49c06.x
c	6	107.6	52.7	336	42	AI085076	AI085076 ow88h01.s
c	7	99	48.5	466	29	AA170068	AA170068 ms44c07.r

No.						
8	76.4	37.5	459	26	W86797	W86797 zh64c05.s1
9	70	34.3	551	34	AA496651	AA496651 zv37t09.r
10	61.8	30.3	551	34	AW102885	AW102885 xd39g11.x
11	61.6	30.2	433	69	AW757285	AW757285 ah69f09.s
12	61.6	30.2	454	38	AA865734	AA865734
13	61.6	30.2	427	39	AI355317	AI355317 oh41g09.x
14	61.6	28.6	418	45	AI184462	AI184462
15	58.4	28.5	263	29	AA203802	AA203802 mu61b08.r
16	47.6	28.0	440	24	H83744	H83744 yv84d03.r1
17	46	23.3	443	24	AA448534	AA448534 zw81e03.r
18	42	22.5	322	33	AI921056	AI921056 w018h07.x
19	41.8	20.6	715	62	AQ487699	AQ487699 RPCI-11-2
20	41.2	20.5	761	103	AI311063	AI311063 uj25b07.x
21	39.6	20.2	502	44	AI581344	AI581344 ta31h11.x
22	39.2	19.2	486	41	AI015911	AI015911 ov26a06.x
23	39	19.1	531	41	AV252285	AV252285
24	38.6	19.1	238	70	AQ249576	AQ249576 F16C24-T7
25	38.2	18.7	663	99	C90865	C90865 Dict
26	37.4	18.3	437	40	AU011742	AU011742 RPCI-11-1
27	37	18.1	357	102	AQ415510	AQ415510 RPCI-11-1
30	36.8	18.0	1101	79	CNS008XS	AL052569 Drosophil
31	36.8	18.0	896	81	B10661	B10661 F16C24-T7
32	36.6	17.9	319	40	C92584	C92584 C92584 Dict
33	36.6	17.9	565	40	C93739	C93739 C93739 Dict
34	36.4	17.8	585	102	AQ384109	AQ384109 RPCI11-15
35	36.2	17.7	1101	79	CNS008GV	AL051859 Drosophil
36	36.2	17.7	1117	79	CNS0022S	AL098338 Drosophil
36	36	17.6	430	100	AQ254560	AQ295460 HS_3048_B
37	35.8	17.6	450	22	H15281	H15281 ym30f10.s1
38	35.8	17.5	486	61	AL041544	AL041544 DKFZp434H
39	35.6	17.5	1101	79	CNS017KX	AL108171 Drosophil
40	35.6	17.5	383	23	H59015	H59015 yr40a11.s1
40	35.6	17.5	566	99	AQ239014	AQ239014 RPCI11-70
41	35.4	17.4	559	40	C89995	C89995 C89995 Dict
42	35.4	17.4	240	51	AU073371	AU073371 AU073371
43	35.4	17.4	240	51	AU073690	AU073690 AU073690
44	35.2	17.3	240	51	AG020469	AG020469 Homo sapi
45	35.2	17.3	1101	79	CNS006EA	AL065155 Drosophil

ALIGNMENTS

Query Match   90.3% ;   Score 184.2; DB 74;   Length 515;
Best Local Similarity 98.0%;   Pred. No. 7.4e-34;
Matches 197; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```
Qy      1 gatcgttgtacctcttcaagcaaaatgaaaaattcttcatcataatttttgccaaaacttcga   60
          ||||| |||||||||||||||||||||||||||||| ||||| ||||||||||||||||||||
Db    315 GATCCTTGTACCTCTTCAAGCAAAATGAAAAATTCTTTCATAATTTTTGCCAAAACTTCGA  374

Qy     61 acttcttcattaattgataatttatatcatcatcataatttcttcttgccatggccaggagat  120
          ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    375 ACTTCTTCATTAATTGATAATTTATATCATCATCATAATTTCTTCTTGCCATGGCCAGGAGAT  434

Qy    121 atttgttgtgtatttggcacagtt-cacggaggacacttactggttgggtgggtcacgt  179
          ||||||||||||||||||||||||| ||  | ||||| ||||| | |||| |||||||| |
Db    435 ATTTGTTGTGTATTTGGCACAGTTTCCACCAAAGACATTACTGTGTAACTGTGTGGGTCACT  494

Qy    180 gcaggagacactgtagtttttg  200
          |||| ||||||||||||||||||
Db    495 GCAGAGAGACACTGTAGTTTTTG  515
```

found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Seq primer: -40UP from Gibco
High quality sequence stop: 454.

FEATURES   Location/Qualifiers
     source   1..510
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2479741"
                     /clone_lib="NCI_CGAP_GC6"
                     /tissue_type="pooled germ cell tumors"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Plasmid DNA from the normalized library
                     NCI_CGAP_GC4 was prepared, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from a pool of 5,000 clones made
                     from the same library (cloneIDs 1257096-1258631,
                     1469064-1470983, and 1475592-1476743). Subtraction by
                     Bento Soares and M. Fatima Bonaldo."
BASE COUNT   131 a   105 c   82 g   192 t
ORIGIN

Query Match   87.8%;   Score 179.2; DB 63; Length 510;
Best Local Similarity   98.0%;   Pred. No. 1.1e-32;
Matches 192; Conservative   0; Mismatches   3; Indels   1; Gaps   1;

Qy   1 gatcgttgtacctcctcaagcaaaatgaaaattcttcataattttgcccaaacctcga 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   315 GATCGTTGTACCTCCTCAAGCAAAATGAAAATTCTTCATAATTTTGCCCAAACCTTCGA 374

Qy   61 acttctttcattaattgatatttttatatcatcatcattaattctttcttgcccatggcagatt 120
          ||||| ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db   375 ACTTCTTTCATTAATTGATATTTTTATATCATCATCATTAATTCTTTCTTGCCCATGGCAGGAGAT 434

Qy   121 atttgtgttggtgtatttggcacagttt--caccgagacaacattactgtaactgtggtgggtcacgt 179
          ||||| |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   435 ATTTGTGGTGTATTTGGCACAGTTTCCACCAAAAGACATTACTGTAACTGGTGGGTCACCT 494

Qy   180 gcagagacactgtagt 195
          ||||||||||||||||
Db   495 GCAGAGACACTGTAGT 510

DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

LOCUS   AI624830   505 bp   mRNA   EST   22-APR-1999
DEFINITION ts71fi2.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2235751 3'
            similar to TR:Q61204 Q61204 NOTCH2-LIKE ;. mRNA sequence.
ACCESSION   AI624830
VERSION   AI624830.1 GI:4649761
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 505)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
            Unpublished (1997)
COMMENT   On Jun 5, 1998 this sequence version replaced gi:3187074.
  Contact: Robert Strausberg, Ph.D.
  Email: Robert_Strausberg@nih.gov
  Tel: (301) 496-1550
  Tissue Procurement: Robert A. Moskaluk, M.D., Ph.D., Michael
  R. Emmert-Buck, M.D., Ph.D.
  cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
  Bonaldo, Ph.D.
  cDNA Library Arrayed by: Greg Lennon, Ph.D.

Seq primer: -40UP from Gibco
High quality sequence stop: 452.

FEATURES   Location/Qualifiers
     source   1..505
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2236751"
                     /clone_lib="NCI_CGAP_GC6"
                     /tissue_type="pooled germ cell tumors"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Plasmid DNA from the normalized library
                     NCI_CGAP_GC4 was prepared, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from a pool of 5,000 clones made
                     from the same library (cloneIDs 1257096-1258631,
                     1469064-1470983, and 1475592-1476743). Subtraction by
                     Bento Soares and M. Fatima Bonaldo."
BASE COUNT   130 a   105 c   80 g   190 t
ORIGIN

Query Match   85.4%;   Score 174.2; DB 49; Length 505;
Best Local Similarity   97.9%;   Pred. No. 1.6e-31;
Matches 187; Conservative   0; Mismatches   3; Indels   1; Gaps   1;

Qy   1 gatcgttgtacctcctcaagcaaaatgaaaattcttcataattttgcccaaacctcga 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   315 GATCGTTGTACCTCCTCAAGCAAAATGAAAATTCTTCATAATTTTGCCCAAACCTTCGA 374

Qy   61 acttctttcattaattgatatttttatatcatcatcattaattctttcttgcccatggcagatt 120
          ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   375 ACTTCTTTCATTAATTGATATTTTTATATCATCATCATTAATTCTTTCTTGCCCATGGCAGGAGAT 434

Qy   121 atttgtgttggtgtatttggcacagttt--caccgagacaacattactgtaactgtggtgggtcacgt 179
          ||||| |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   435 ATTTGTGGTGTATTTGGCACAGTTTCCACCAAAAGACATTACTGTAACTGGTGGGTCACCT 494

Qy   180 gcagagacact 190
          |||||||||||
Db   495 GCAGAGACACT 505

RESULT   4
AA155054/c
AA155054
LOCUS   AA155054   330 bp   mRNA   EST   11-FEB-1997
DEFINITION mr91bl1.r1 Stratagene mouse embryonic carcinoma (#937317) Mus
            musculus cDNA clone IMAGE:604797 5' similar to TR:G1336628 G1336628
            EGF REPEAT TRANSMEMBRANE PROTEIN. ;. mRNA sequence.
ACCESSION   AA155054
VERSION   AA155054.1 GI:1726674
KEYWORDS   EST.
SOURCE   house mouse.
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 330)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE   The WashU-HHMI Mouse EST Project
            Unpublished (1996)
COMMENT   On Sep 12, 1996 this sequence version replaced gi:1292301.
  Contact: Marra M/Mouse EST Project
  WashU-HHMI Mouse EST Project

Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mousees@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Possible reversed clone: similarity on wrong strand
Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 85.

FEATURES
source   1..330
                /organism="Mus musculus"
                /db_xref="taxon:10090"
                /clone="IMAGE:604797"
                /clone_lib="Stratagene mouse embryonic carcinoma (#937317)"
                /tissue_type="carcinoma"
                /dev_stage="embryonic"
                /lab_host="SOLR (kanamycin resistant)"
                /note="Vector:pBluescript SK+; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. P19 cell line. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCACGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTTTTT 3'"

BASE COUNT   113 a   62 c   72 g   83 t
ORIGIN

Query Match   53.9%;   Score 110; DB 29; Length 330;
Best Local Similarity 74.8%; Pred. No. 1.8e-16;
Matches 151; Conservative 0; Mismatches 50; Indels 1; Gaps 1;

Qy   1 gatcgttgtacctcttcaagcaaaatgaaaaattcttcataaattcttgcccaaaccttcga 60
        |||||| ||||||||| ||||||||||| |||||||||| ||||||||| ||||||| |||
Db 272 GGTCCGTGCGCACTTCTTCAAGCAAAATGAAAAATCCTTCAACTTCGTCCAAAACTTCGA 213

Qy   61 acttcttcattaattgatatttttatcatcattaatttctttgccatggcaggagat 120
        |||||||| |||| | || |||||| |||||||||| ||| ||||||| ||| |||
Db 212 ACTTCCTTCATTAATTGATGCTTTATGTCACCATTGATGCCTTGACCCATTGCAGGAGTG 153

Qy 121 atttgtggtgtattggcacagttt-caccgaagacattactgtaactggtggtcacgt 179
        ||||| ||| |||||||||||||| |||||||||||| |||| ||| |||||||| |
Db 152 ATTTGCATTGAATTTGGCACCGTTCCAAACCAGATGCCATTGTGATTTTTGGTCAGCT 93

Qy 180 gcaggagacactgtagttttttgg 201
        |||| ||||||| |||| ||||||
Db   92 ACAGAAGAGACTGCAGGTTTTGG 71

RESULT 5
LOCUS   AI654710   425 bp   mRNA   EST   04-MAY-1999
DEFINITION wb49c06.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2309002 3' similar to TR:O61204 O61204 NOTCH2-LIKE ;. mRNA sequence.
ACCESSION   AI654710
VERSION   AI654710.1 GI:4738689
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 425)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   On May 18, 1998 this sequence version replaced gi:3138277.
            Contact: Robert Strausberg, Ph.D.
            Email: Robert_Strausberg@nih.gov
            Tel: (301) 496-1550
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael

R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Seq primer: -40UP from Gibco
High quality sequence stop: 404.

FEATURES
source   1..425
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:2309002"
                /clone_lib="NCI_CGAP_GC6"
                /tissue_type="pooled germ cell tumors"
                /lab_host="DH10B"
                /note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI. NCI_CGAP_GC4 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIds 1257096-1258631, 1469064-1470983, and 1475592-1476743). Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT   109 a   89 c   59 g   167 t   1 others
ORIGIN

Query Match   53.1%;   Score 108.4; DB 49; Length 425;
Best Local Similarity 98.2%; Pred. No. 4e-16;
Matches 109; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy   1 gatcgttgtacctcttcaagcaaaatgaaaaattcttcataaattcttgcccaaaccttcga 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db 315 GATCGTTGTACCTCTTCAAGCAAAATGAAAAATTCTTCATAAATTCTTTGCCCAAAACTTCGA 374

Qy   61 acttcttcattaattgatatttttatcatcattaatttctttgccatg 111
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db 375 ACTTCTTCATTAATTGATATTNTATCATCATTAATTTCTTTGCCATG 425

RESULT 6
LOCUS   AI085076   336 bp   mRNA   EST   28-AUG-1998
DEFINITION ow88h01.s1 Soares_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA clone IMAGE:1653937 3' similar to TR:O61204 O61204 EGF REPEAT TRANSMEMBRANE PROTEIN. ;. mRNA sequence.
ACCESSION   AI085076
VERSION   AI085076.1 GI:3423499
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 336)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   On Jan 19, 1998 this sequence version replaced gi:2285511.
            Clone IMAGE:1653937 3' similar to TR:O61204 O61204 EGF REPEAT
            Contact: Robert Strausberg, Ph.D.
            Email: Robert_Strausberg@nih.gov
            Tel: (301) 496-1550
            Insert Length: 790 Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 334.

FEATURES

RESULT 7
AA170068/c

LOCUS   AA170068   466 bp   mRNA   EST   20-DEC-1996
DEFINITION ms44c07.r1 Life Tech mouse embryo 13 5dpc 10656014 Mus musculus
            cDNA clone IMAGE:614412 5' similar to TR:G1336628 G1336628 EGF
            REPEAT TRANSMEMBRANE PROTEIN. ;, mRNA sequence.
ACCESSION   AA170068
VERSION   AA170068.1   GI:1748620
KEYWORDS   EST.
SOURCE   house mouse.
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 466)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE   The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
  COMMENT   On Sep 12, 1996 this sequence version replaced gi:1405046.
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:375236
            Possible reversed clone: similarity on wrong strand

Query Match   52.7%; Score 107.6; DB 42; Length 336;
Best Local Similarity   83.6%; Pred. No. 6.4e-16;
Matches 122; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy   1 gatcgttgtaccctcttcaagcaaaatgaaaattcttcatataatttgcccaaaccttcga 60
          |||-||||||||||||||||||||||||||||||| |||||||||||||||||| |||||||
Db   177 GGTCCTTGCACTTCTTCAAGCAAAATGAAAATTCTTCATATATTTCGACCAAACTTTCGA 236

Qy   61 acttcttcattaaattagatatattttatcatcatcattaatttctttggccattggcaggagat 120
          |||| ||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db   237 ACTTCCTTCATTAATTGATGGTTTTATATCAGCATTGAATCCTTGTGCCATAGCAGGAGTG 296

Qy   121 atttgtggttgttattggcacagttc 146
          |||-||||||||||||||||||||||
Db   297 ATTTGTAAGTGTATTTGGCATAGCTCC 322

RESULT 8
W86797

LOCUS   W86797   459 bp   mRNA   EST   01-JUL-1996
DEFINITION zh64c05.s1 Soares_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA
            clone IMAGE:416840 3', mRNA sequence.
ACCESSION   W86797
VERSION   W86797.1   GI:1400525
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 459)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE   The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
  COMMENT   On May 16, 1995 this sequence version replaced gi:808442.
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: mob.REGA+ET
            High quality sequence stop: 361.

FEATURES   Location/Qualifiers
     source   1. .459

Query Match   48.5%; Score 99; DB 29; Length 466;
Best Local Similarity   74.4%; Pred. No. 6.1e-14;
Matches 151; Conservative 0; Mismatches 50; Indels 2; Gaps 2;

Qy   1 gatcgttgtaccctcttcaagcaaaatgaaaattcttcatataatttgcccaaaccttcga 60
          |-|| ||||||||||||||||||||||||||||||||||||||||| |||||| |||||||
Db   393 GGTCCGTCGCACTTCTTCAAGCAAAATGAAAATTCTTCATATATTTCTTTTCATATACTTTCGA 334

Qy   61 acttcttcattaaattagatatattttatcatcatcattaatttctttggccattggcaggagat 120
          ||||||||||||||||||||||| |||||||||||||||||||||||| |||||||||||||
Db   333 ACTTCCTTCATTAATTGATGGTTTTATATCAGCATTGATGCCTTGAGCCATTGCAGGAGTG 274

Qy   121 atttgtggttgttattggcacagttt-caccgaagacattactgt-aactggtgggtcacg 178
          |||-|||||||||||||||||||| || ||||| ||| || |||| | |||||| ||||||
Db   273 ATTTGCATTGAATTGGCACCGTTTCCAAAGCACAGATGCCATTGTGCATTTTTTTCGGTCAGC 214

Qy   179 tgcaggagacactgtagttctttttg 201
          ||-|||| ||| |||| |||||||
Db   213 TACAGAAAGACTGCAGGTTTTGG 191

---

source
          1. .336
          /organism="Homo sapiens"
          /db_xref="taxon:9606"
          /clone="IMAGE:1653937"
          /clone_lib="Soares_fetal_liver_spleen_1NFLS_S1"
          /sex="male"
          /dev_stage="20 week-post conception fetus"
          /lab_host="DH10B (ampicillin resistant)"
          /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
          with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
          This is a subtracted version of the original Soares fetal
          liver spleen 1NFLS library. The library was primed
          with a Pac I - oligo(dT) primer [5'
          AACTGGAAGAATTAATTAAGATCTTTTTTTTTTTTTTT 3'],
          double-stranded cDNA was ligated to Eco RI adaptors
          (Pharmacia), digested with Pac I and cloned into the Pac I
          and Eco RI sites of the modified pT7T3 vector. Library
          went through one round of normalization. Library
          constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT   84 a   64 c   62 g   126 t
ORIGIN

FEATURES   Location/Qualifiers
     source   1. .466
                     Seq primer: -28M13 rev1 from Amersham
                     High quality sequence stop: 345.
          /organism="Mus musculus"
          /strain="C57BL/6J"
          /db_xref="taxon:10090"
          /clone="IMAGE:614412"
          /clone_lib="Life Tech mouse embryo 13 5dpc 10656014"
          /note="Organ: whole embryo; Vector: pCMV-SPORT2; Site_1:
          Sal1; Site_2: Not1; Cloned unidirectionally. Primer:
          Oligo dT. 13.5dpc embryos. pCMV-SPORT2 vector."
          /dev_stage="13.5dpc embryos"
          /tissue_type="embryo"
          /lab_host="DH10B"

BASE COUNT   157 a   82 c   104 g   122 t   1 others
ORIGIN

RESULT 9
LOCUS   AA496651/c
DEFINITION AA496651   551 bp   mRNA   EST   12-AUG-1997
            zv37f09.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone
            IMAGE:755849 5' similar to TR:G1336628 G1336628 EGF REPEAT
            TRANSMEMBRANE PROTEIN. ;, mRNA sequence.
ACCESSION   AA496651
VERSION   AA496651.1   GI:2229972
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 551)
AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,
            Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE   WashU-Merck EST Project 1997
JOURNAL   Unpublished (1997)
COMMENT   On Jan 25, 1995 this sequence version replaced gi:637829.
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 508.
FEATURES   Location/Qualifiers
     source   1..551
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:755849"
                     /clone_lib="Soares ovary tumor NbHOT"

Query Match   37.5%;   Score 76.4;   DB 26;   Length 459;
Best Local Similarity   84.3%;   Pred. No. 1.2e-08;
Matches   86;   Conservative   0;   Mismatches 16;   Indels   0;   Gaps   0;

Qy   1 gatcgttgtacctcttcaagcaagacaaaatgaaaaattcttcaagtgcccaaaccttcga 60
               |||||||||||||||||||||||||||||||||||||| |||||||||| |||||||||
Db   358 GGTCCTTGCACTTCTTCAAGCAAAATGAAAAATTCTTTCAATTTCGACCAAACTTTCGA 417

Qy   61 acttcttcattaattgatatttttatatcatcattaatttc 102
               ||||||||||||| ||||||||||||||||||| ||||| |
Db   418 ACTTCCTTCATTAATTGATGGTTTATATCAGCATTGGATTCC 459

RESULT 10
LOCUS   AW102885   433 bp   mRNA   EST   19-OCT-1999
DEFINITION AW102885
            xd39g11.x1 NCI_CGAP_Ov23 Homo sapiens cDNA clone IMAGE:2596196 3'
            similar to TR:Q61204 Q61204 NOTCH2-LIKE ;, mRNA sequence.
ACCESSION   AW102885
VERSION   AW102885.1   GI:6073498
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 433)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   On Jun 22, 1998 this sequence version replaced gi:3246955.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Possible reversed clone: polyT not found
            Seq primer: -40UP from Gibco
            High quality sequence stop: 358.
FEATURES   Location/Qualifiers
     source   1..433
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2596196"
                     /clone_lib="NCI_CGAP_Ov23"
                     /tissue_type="tumor, 5 pooled (see description)"
                     /lab_host="DH10B"
                     /note="Organ: ovary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.35 kb. Tumor types include: mixed

Query Match   34.3%;   Score 70;   DB 34;   Length 551;
Best Local Similarity   88.4%;   Pred. No. 3.6e-07;
Matches   76;   Conservative   0;   Mismatches 10;   Indels   0;   Gaps   0;

Qy   1 gatcgttgtacctcttcaagcaagacaaaatgaaaaattcttcaagtgcccaaaccttcga 60
               ||||||||||||||||||||||||||||||||||||| ||||||||| |||||||||
Db   118 GGTCCTTGCACTTCTTCAAGCAAAATGAAAAATTCTTTCAATTTCGACCAAACTTTCGA 59

Qy   61 acttcttcattaattgatatttttat 86
               ||||||||||||||||||||||| ||
Db   58 ACTTCCTTCATTAATTGATGTTTTAT 33

BASE COUNT   180 a   89 c   121 g   161 t
ORIGIN

        /sex="female"
        /tissue_type="ovarian tumor"
        /lab_host="DH10B (ampicillin resistant)"
        /note="Organ: ovary; Vector: pT7T3D (Pharmacia) with a
        modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
        strand cDNA was primed with a Not I - oligo(dT) primer [5'
        TGTTACCAATCTGAAGTGGGAGCGCGCCGCCGCGGTTTTTTTTTTTTTTT 3'], 1st
        strand cDNA was size selected, ligated to Eco RI
        adapters (Pharmacia), digested with Not I and cloned into
        the Not I and Eco RI sites of a modified pT7T3 vector
        (Pharmacia). Library constructed by Bento Soares and
        M.Fatima Bonaldo."

BASE COUNT   147 a   89 c   63 g   160 t
OF

        /organism="Homo sapiens"
        /db_xref="GDB:132531O"
        /db_xref="taxon:9606"
        /clone="IMAGE:416840"
        /sex="male"
        /clone_lib="Soares_fetal_liver_spleen_1NFLS_S1"
        /dev_stage="20 week-post conception fetus"
        /lab_host="DH10B (ampicillin resistant)"
        /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
        with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
        This is a subtracted version of the original Soares fetal
        liver spleen 1NFLS library.   1st strand cDNA was primed
        with a Pac I - oligo(dT) primer [5'
        AACTGGAAGAATTAAAGAGATCTTTTTTTTTTTTTTTTTT 3'],
        double-stranded cDNA was ligated to Eco RI adaptors
        (Pharmacia), digested with Pac I and cloned into the Pac I
        and Eco RI sites of the modified pT7T3 vector. Library
        went through one round of normalization. Library
        constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT
ORIGIN   109 a   90 c   71 g   163 t

Mullerian tumor, papillary serous, clear cell, spindle
cell. All are primary tumors, metastasis positive. Life
Technologies catalog #: 11534-013"

Qy   Query Match   30.3%;   Score 61.8;   DB 69;   Length 433;
Best Local Similarity   74.3%;   Pred. No. 3.1e-05;
Matches   78;   Conservative   0;   Mismatches   27;   Indels   0;   Gaps   0;

Qy   1 gatcgttgtacctcttcaagcaaaatgaaaattcttcattaatttgcccaaaccttcga 60
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   321 GGTCCTTGCACTCCTTCAAGCATTTCGAAGATTTTTTCATATTTTTGTCAAGGCATCGG 380

Qy   61 acttctttcattaattgatcattttatatccatcattccattcttt 105
       | ||| |||| || |||||||||| |||||||| || |||||||| | |
Db   381 ATTTTCCTTCACTACTTGACATTTATTATATCAGCATTAATTCTTGT 425

RESULT 11
LOCUS   AA757285   454 bp   mRNA   EST   29-DEC-1998
DEFINITION   ah69f09.s1 Soares_testis_NHT Homo sapiens cDNA clone 1320905 3',
            mRNA sequence.
ACCESSION   AA757285
VERSION   AA757285.1   GI:2805148
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 454)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   On Jan 19, 1998 this sequence version replaced gi:2151913.
            Contact: Robert Strausberg, Ph.D.
            Email: Robert_Strausberg@nih.gov
            Tel: (301) 496-1550
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

FEATURES   Location/Qualifiers
            source   1..454
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="1320905"
                        /clone_lib="Soares_testis_NHT"
                        /sex="male"
                        /lab_host="DH10B"
                        /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                        polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                        was prepared from mRNA obtained from Clontech
                        Laboratories, Inc., and primed with a Not I - oligo(dT)
                        primer [5'
                        TGTTACCAATCTGAAGTGGGAGGGGCGCCGCCCCAATTTTTTTTTTTTTTT 3'].
                        Double-stranded cDNA was ligated to Eco RI adaptors
                        (Pharmacia), digested with Not I and cloned into the Not I
                        and Eco RI sites of the modified pT7T3 vector. Library
                        went through one round of normalization to Cot5, and was
                        constructed by Bento Soares and M. Fatima Bonaldo."
            Insert Length: 320   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham.

BASE COUNT
ORIGIN   119 a   96 c   79 g   160 t

Qy   Query Match   30.3%;   Score 61.6;   DB 38;   Length 454;
Best Local Similarity   76.0%;   Pred. No. 3.4e-05;
Matches   76;   Conservative   0;   Mismatches   24;   Indels   0;   Gaps   0;

Qy   1 gatcgttgtacctcttcaagcaaaatgaaaattcttcattaatttgcccaaaccttcga 60
       ||| | ||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db   324 GGTCCTTGCACTCCTTCAAGCATTTCGAAGATTTTTTCATATTTTTGTCAAGGCATCGG 383

Qy   61 acttctttcattaattgatcattttatatccatcattccattcttt 100
       | ||| |||| || ||||||||||| |||||||| || |||||||| |
Db   384 AGTTCCTTCACTACTTGACATTTATTATATCAGCATTAATTT 423

RESULT 12
LOCUS   AA865734   427 bp   mRNA   EST   13-MAY-1998
DEFINITION   oh41g09.s1 NCI_CGAP_GC4 Homo sapiens cDNA clone IMAGE:1469248 3'
            similar to TR:Q61204 Q61204 EGF REPEAT TRANSMEMBRANE PROTEIN. ;,
            mRNA sequence.
ACCESSION   AA865734
VERSION   AA865734.1   GI:2958010
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 427)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   On Jan 14, 1998 this sequence version replaced gi:1796968.
            Contact: Robert Strausberg, Ph.D.
            Email: Robert_Strausberg@nih.gov
            Tel: (301) 496-1550
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            High quality sequence stop: 315.

FEATURES   Location/Qualifiers
            source   1..427
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:1469248"
                        /clone_lib="NCI_CGAP_GC4"
                        /tissue_type="pooled germ cell tumors"
                        /lab_host="DH10B"
                        /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                        polylinker; 1st strand cDNA was prepared from 3 pooled
                        germ cell tumors, and was then primed with a Not I -
                        oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                        RI adaptors (Pharmacia), digested with Not I and cloned
                        into the Not I and Eco RI sites of the modified pT7T3
                        vector. Library is normalized. Library was constructed by
                        Bento Soares and M. Fatima Bonaldo."
            Insert Length: 874   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham.

BASE COUNT
ORIGIN   109 a   91 c   68 g   159 t

Qy   Query Match   30.3%;   Score 61.6;   DB 39;   Length 427;
Best Local Similarity   76.0%;   Pred. No. 3.4e-05;
Matches   76;   Conservative   0;   Mismatches   24;   Indels   0;   Gaps   0;

Qy   1 gatcgttgtacctcttcaagcaaaatgaaaattcttcattaatttgcccaaaccttcga 60

RESULT 13

```
Db   315 GGTCCCTTGCACTCCTTCAAGCATTCGAAGATTTTTCGAAGATTTTTCATATAATTTTTGTCCAACGCATCGG 374
         |  | |||  |||  ||||||||||   ||||   |||| |||| ||   |  ||||| | |||
Qy    61 acttcttcattaattgatattttatatcatcattaatt 100

Db   375 AGTTCCTTCACTAATTTACGTTTTTATATCAGCATTAATTT 414
         |||| |||||||||||| |||||||  |||||||||||||
Qy       
```

LOCUS   AI355317   418 bp   mRNA   EST   15-FEB-1999
DEFINITION qt57g03.x1 Soares_fetal_lung_NbHH19W Homo sapiens cDNA clone
            IMAGE:1952116 3' similar to TR:Q61204 Q61204 EGF REPEAT
            TRANSMEMBRANE PROTEIN. ;, mRNA sequence.
ACCESSION   AI355317
VERSION   AI355317.1 GI:4095470
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 418)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   On Jan 19, 1998 this sequence version replaced gi:2152926.
            Contact: Robert Strausberg, Ph.D.
            Email: Robert_Strausberg@nih.gov
            Tel: (301) 496-1550
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 552   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 417.
FEATURES   Location/Qualifiers
     source   1..418
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1952116"
                     /clone_lib="Soares_fetal_lung_NbHH19W"
                     /dev_stage="19 weeks"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: lung; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer
                     [5'-TGTTACCAATCTGAAGTGGGAGGCGGCCGCAATTTTTTTTTTTTTTTTT-3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by Bento
                     Soares and M.Fatima Bonaldo. This library was constructed
                     from the same fetus as the fetal heart library, Soares
                     fetal heart NbHH19W."
BASE COUNT   110 a   91 c   73 g   144 t
ORIGIN

```
Query Match             30.2%;   Score 61.6;  DB 45;  Length 418;
Best Local Similarity   76.08%;  
Matches   76;   Conservative    0;   Mismatches   24;   Indels    0;   Gaps    0;

Qy     1 gatcgtgtacctcttcaagcaaaatgaaaattcttcataatttgcccaaacccttcga 60
         ||| ||||||||||||||||| ||||  |||||||||||| || ||||||||||||||
Db   301 GGTCCCTTGCACTCCTTCAAGCANTTCGAAGATTTTTCATATAATTTTTGTCCAACGCATCGG 360

Qy    61 acttcttcattaattgatattttatatcatcattaatt 100
         ||||||||||||| |||||||||  |||||||||||||
Db   361 AGTTCCTTCACTAATTTACGTTTTTATATCAGCATTAATT 400
```

RESULT 14

LOCUS   AA189462   263 bp   mRNA   EST   13-JAN-1997
DEFINITION mt79f05.r1 Soares mouse lymph node NbMLN Mus musculus cDNA clone
            IMAGE:636129 5' similar to TR:G1336628 G1336628 EGF REPEAT
            TRANSMEMBRANE PROTEIN. ;, mRNA sequence.
ACCESSION   AA189462
VERSION   AA189462.1 GI:1777043
KEYWORDS   EST.
SOURCE   house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 263)
AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
TITLE   The WashU-HHMI Mouse EST Project
JOURNAL   Unpublished (1996)
COMMENT   On Sep 12, 1996 this sequence version replaced gi:1406868.
            Contact: Marra M/Mouse EST Project
            Washu-HHMI Mouse EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:388121   -28M13 rev2 from Amersham
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 245.
FEATURES   Location/Qualifiers
     source   1..263
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:636129"
                     /clone_lib="Soares mouse lymph node NbMLN"
                     /sex="male"
                     /tissue_type="lymph node"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
                     /note="Organ: lymph node; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Not I - oligo(dT) primer
                     [5'-
                     TGTTACCAATCTGAAGTGGGAGGCGGCCGCGATACTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with the modified pT7T3 vector. RNA
                     and Eco RI sites of the modified pT7T3 vector, RNA
                     provided by Dr. Bertrand Jordan. Library constructed and
                     normalized by Dr. Bento Soares and M.Fatima Bonaldo."
BASE COUNT   100 a   42 c   56 g   65 t
ORIGIN

```
Query Match             28.6%;   Score 58.4;  DB 29;  Length 263;
Best Local Similarity   85.5%;   Pred. No. 0.00021;
Matches   65;   Conservative    0;   Mismatches   11;   Indels    0;   Gaps    0;

Qy     1 gatcgtgtacctcttcaagcaaaatgaaaattcttcataatttgcccaaacccttcga 60
         ||| ||||||||||||||||| |||||||||||||||| |||| ||||||||||||||
Db    76 GGTCCCCTGCACTTCTTCAAGCANAAATGAAAATCCTTTCATACTTTTCGTCCAAACTTTCGA 17

Qy    61 acttcttcattaatt 76
         |||||||||||||||
Db    16 ACTTCCTTCATTAATT 1
```

RESULT 15
AA203802/c

```
LOCUS       AA203802      440 bp    mRNA                      EST       27-JAN-1997
DEFINITION  mu61a08.r1 Soares mouse lymph node NbMLN Mus musculus cDNA clone
            IMAGE:643862 5' similar to TR:G1336628 G1336628 EGF REPEAT
            TRANSMEMBRANE PROTEIN: :, mRNA sequence.
ACCESSION   AA203802
VERSION     AA203802.1  GI:1800348
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 440)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     On May 9, 1995 this sequence version replaced gi:803991.
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:395854
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 431.
FEATURES             Location/Qualifiers
     source          1..440
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:643862"
                     /clone_lib="Soares mouse lymph node NbMLN"
                     /sex="male"
                     /tissue_type="lymph node"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
                     /note="Organ: lymph node; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Not I - oligo(dT) primer
                     [5'
                     TGTTACCAATCTGAAGTGGGAGGGCCGCGAPACTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. RNA
                     provided by Dr. Bertrand Jordan. Library constructed and
                     normalized by Bento Soares and M.Fatima Bonaldo."
BASE COUNT      144 a     87 c    100 g    109 t
ORIGIN
```

```
Query Match        28.6%;     Score 58.4.;  DB 30;    Length 440;
Best Local Similarity  85.5%;   Pred. No. 0.00019;
Matches   65;  Conservative    0;  Mismatches   11;  Indels   0;  Gaps    0;

Qy         1 gacgttgtaccctcaagcaaaatgaaaatctttcaataatttgcccaaaccttga  60
             |||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        76 GGTCCCTGCACTTCTTCAAGCAAAATGAAAATCCTTTCATAACTTTCGTCCAAACTTTCGA 17

Qy        61 acttctttcaattaatt  76
             ||||| ||||||||||||
Db        16 ACTTCCTTCATTAATT 1
```

Search completed: March 23, 2000, 21:51:26
Job time: 8110 sec

⊔⊃.   ⊥⊐ TACAԌ⌐⌐⌐.ԌACTGCAGGTTT⊥⌐⌐ ⊥91

RESULT   8
W86797
LOCUS   W86797   459 bp   mRNA   EST   01-JUL-1996
DEFINITION   zh64c05.s1 Soares_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA
               clone IMAGE:416840 3', mRNA sequence.
ACCESSION   W86797
VERSION   W86797.1 GI:1400525
KEYWORDS   EST.
SOURCE   human.
  ORGANISM   Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
               Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 459)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
               Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
               Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
               Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
               Wilson,R.
  TITLE   The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT   On May 16, 1995 this sequence version replaced gi:808442.
               Contact: Wilson RK
               Washington University School of Medicine
               4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
               Tel: 314 286 1800
               Fax: 314 286 1810
               Email: est@watson.wustl.edu
               This clone is available royalty-free through LLNL ; contact the
               IMAGE Consortium (info@image.llnl.gov) for further information.
               Seq primer: mob.REGA+ET
               High quality sequence stop: 361.
FEATURES   Location/Qualifiers
     source   1..459
                     /organism="Homo sapiens"
                     /db_xref="GDB:1325310"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:416840"
                     /clone_lib="Soares_fetal_liver_spleen_1NFLS_S1"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     This is a subtracted version of the original Soares fetal
                     liver spleen 1NFLS library. 1st strand cDNA was primed
                     with a Pac I - oligo(dT) primer [5'
                     AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT   147 a   89 c   63 g   160 t
ORIGIN

  Query Match   37.5%; Score 76.4; DB 26; Length 459;
  Best Local Similarity   84.3%; Pred. No. 1.2e-08;
  Matches   86; Conservative   0; Mismatches   16; Indels   0; Gaps   0;

Qy   1 gatcgttgtacctcttcaagcaaaatgaaaattctttcataatttttgcccaaaccttcga 60
          | || ||| || ||||||||||||||||||||||||||||||||||| ||| | |||||| |||||
Db   358 GGTCCTTGCACTTCTTCAAGCAAAATGAAAATTCTTTCATATTTTCGACCAAACTTTCGA 417

Qy   61 acttctttcattaattgatattttatatcatcattaattttc 102
          ||||| |||||||||||||| ||||||||||| |||| || |
Db   418 ACTTCCTTCATTAATTGATGGTTTATATCAGCATTGGATTCC 459

RESULT   9
AA496651/c
LOCUS

THIS PAGE BLANK (USPTO)

OM nucleic - nucleic search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on:　　　March 24, 2000, 13:21:14 ; Search time 172.6 Seconds
　　　　　　　　　　　　　　　　　　　(without alignments)
　　　　　　　　　　　　　　　　　　　1708.448 Million cell updates/sec

Title:　　　　US-09-183-789-43
Perfect score:　2463
Sequence:　　1 gcttggtaccgagctcggat..........aaaaaaaaaaaaaaaa 2463

Scoring table:　IDENTITY_NUC
　　　　　　　　Gapop 10.0 , Gapext 1.0

Searched:　　214294 seqs, 59861574 residues

⬤ 1 number of hits satisfying chosen parameters:　428588

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
　　　　　　　　　Listing first 45 summaries

Database :　　Issued_Patents_NA:*
　1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
　2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
　3: /cgn2_6/ptodata/1/ina/5C_COMB.seq:*
　4: /cgn2_6/ptodata/1/ina/5D_COMB.seq:*
　5: /cgn2_6/ptodata/1/ina/6_COMB.seq:*
　6: /cgn2_6/ptodata/1/ina/PCTUS9_COMB.seq:*
　7: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB	ID	Description
1	60.2	2.4	19124	4	US-08-487-826B-13	Sequence 13, Appl
2	56	2.3	665	1	US-08-883-795A-36	Sequence 36, Appl
3	54	2.3	19124	4	US-08-487-826B-13	Sequence 13, Appl
4	53.8	2.2	7218	1	US-08-232-463-14	Sequence 14, Appl
5	52.4	2.2	4525	3	US-08-613-861-2	Sequence 2, Appl
6	51.6	2.1	1123	1	US-08-458-023-3	Sequence 3, Appl
7	50.6	2.1	5852	1	US-07-867-106-2	Sequence 2, Appli
8	50.4	2.0	1511	1	US-07-991-867B-8	Sequence 8, Appli
9	50.4	2.0	1511	2	US-08-107-755A-8	Sequence 8, Appli
10	50.4	2.0	1511	1	US-07-991-867B-8	Sequence 8, Appli
11	50.4	2.0	1511	2	US-08-107-755A-8	Sequence 8, Appli
12	50.2	2.0	9837	3	US-08-544-332-68	Sequence 68, Appl
13	50.2	2.0	9837	3	US-08-832-877-68	Sequence 68, Appl
14	50	2.0	3984	1	US-08-040-753-1	Sequence 1, Appl
15	50-	2.0	4059	6	US-08-252-493C-4	Sequence 4, Appl
16	50	2.0	4059	6	PCT-US95-07554-4	Sequence 4, Appl
17	50	2.0	6420	5	PCT-US96-05511A-11	Sequence 11, Appl
18	50	2.0	9592	1	US-08-393-734-3	Sequence 3, Appl
19	48.6	2.0	665	4	US-08-883-795A-36	Sequence 36, Appl
20	48.4	2.0	1982	4	US-08-900-148-1	Sequence 1, Appl
21	48.2	2.0	660	1	US-07-991-867B-32	Sequence 32, Appl
22	48.2	2.0	660	2	US-08-107-755A-32	Sequence 32, Appl
23	48.2	2.0	660	1	US-08-544-332-32	Sequence 32, Appl
24	48	2.0	404	2	US-08-594-031-154	Sequence 154, App
25	47.8	1.9	1700	4	US-08-839-581A-1	Sequence 1, Appli
26	47.8	1.9	115	3	US-08-454-557C-75	Sequence 75, Appli
27	47.8	1.9	115	4	US-08-340-426D-75	Sequence 75, Appli
28	47.8	1.9	115	4	US-08-450-673C-75	Sequence 75, Appl
29	47.8	1.9	115	6	PCT-US99-17111A-75	Sequence 75, Appl
30	47.8	1.9	5852	1	US-07-867-106-2	Sequence 2, Appli
31	47.6	1.9	6243	4	US-09-056-075-1	Sequence 1, Appli
32	47.2	1.9	2447	4	US-09-014-066-14	Sequence 14, Appli
33	47	1.9	4098	4	US-08-605-106-4	Sequence 4, Appli
34	46.6	1.9	1672	1	US-08-172-331B-13	Sequence 13, Appli
35	46.2	1.9	120	3	US-08-454-557C-76	Sequence 76, Appli
36	46.2	1.9	120	4	US-08-340-426D-76	Sequence 76, Appl
37	46.2	1.9	120	4	US-08-450-673C-76	Sequence 76, Appl
38	46.2	1.9	120	6	PCT-US99-17111A-76	Sequence 76, Appl
39	45.8	1.9	1629	5	PCT-US91-02560-3	Sequence 3, Appli
40	45.8	1.9	4673	1	US-07-638-431-1	Sequence 1, Appli
41	45.8	1.9	4673	6	PCT-US92-00018-1	Sequence 1, Appli
42	45.6	1.9	240	1	US-08-628-417-6	Sequence 6, Appli
43	45.6	1.9	642	2	US-08-764-100-13	Sequence 13, Appli
44	45.6	1.9	643	2	US-08-764-100-7	Sequence 7, Appli
45	45.6	2	2993	2	US-08-764-100-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-487-826B-13/c
Sequence 13, Application US/0848 7826B
Patent No. 5993827
GENERAL INFORMATION:
　APPLICANT: Sim, Kim L.
　APPLICANT: Chitnis, Chetan
　APPLICANT: Miller, Louis H.
　APPLICANT: Peterson, David S.
　APPLICANT: Su, Xin-zhuan
　APPLICANT: Wellems, Thomas E.
　TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
　TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
　NUMBER OF SEQUENCES: 45
　CORRESPONDENCE ADDRESS:
　ADDRESSEE: Knobbe Martens Olson & Bear
　STREET: 620 Newport Center Drive 16th Floor
　CITY: Newport Beach
　STATE: California
　COUNTRY: US
　ZIP: 92660
　COMPUTER READABLE FORM:
　MEDIUM TYPE: Floppy disk
　COMPUTER: IBM PC compatible
　OPERATING SYSTEM: PC-DOS/MS-DOS
　SOFTWARE: PatentIn Release #1.0, Version #1.25
　CURRENT APPLICATION DATA:
　APPLICATION NUMBER: US/08/487,826B
　FILING DATE: 10-SEP-1993
　CLASSIFICATION: 435
　ATTORNEY/AGENT INFORMATION:
　NAME: Israelsen, Ned
　REGISTRATION NUMBER: 29,655
　REFERENCE/DOCKET NUMBER: NIH121.001CP1
　TELECOMMUNICATION INFORMATION:
　TELEPHONE: (619) 235-8550
　TELEFAX: (619) 235-0176
　INFORMATION FOR SEQ ID NO: 13:
　SEQUENCE CHARACTERISTICS:
　LENGTH: 19124 base pairs
　TYPE: nucleic acid
　STRANDEDNESS: single
　TOPOLOGY: linear
　MOLECULE TYPE: cDNA
　HYPOTHETICAL: NO
　ANTI-SENSE: NO
US-08-487-826B-13

Query Match　　　　2.4%;　　Score 60.2;　　DB 4;　　Length 19124;

RESULT 2
US-08-883-795A-36/c
; Sequence 36, Application US/08883795A
; Patent No. 5985607
; GENERAL INFORMATION:
; APPLICANT: Delcuve, Genevieve
; APPLICANT: Awang, Gregor
; TITLE OF INVENTION: Recombinant DNA Molecules and Expression
; TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator
; NUMBER OF SEQUENCES: 39

Best Local Similarity 46.9%; Pred. No. 0.00019;
Matches 222; Conservative 0; Mismatches 248; Indels 3; Gaps 1;

Qy   1928 aatttattacttgaatttgaaactggactgtgttaggagttaaactataattctgt 1987
           |||||||||| ||| |||| || |||||||| ||| | || ||||||||| ||
Db   15966 AATTTTATTTATTTTTTTTTTTTAATTTAAATTTTTTTTTTTTTTTATTCATTTAAT 15907

Qy   1988 tcttaaaagaaaatttatctgcaaatgtgagaatattagctaatgaattagt 2047
           || ||| ||| |||||||||||||| ||| |||| ||| |||||| ||
Db   15906 TTTTTTAATTTTTTTTTTATTTTAATTTTTTTTTTTAATTTATTTATTTAATT 15847

Qy   2048 tgtttggggttactctcttgttctaagatataagaagaatatttgaaactcaa 2107
           || |||||||||| |||||| |||||||||| ||| |||||||||||||||
Db   15846 AAAATTTTTTTATTTTATTTTTTTTTTTAAATAAAATTTTTTTATTTTTGTATATA 15787

Qy   2108 tgaaataattctcagctgcccaaatgtgcacctctttatatatcttttccactttga 2167
           ||||| ||||| ||| |||||||||| ||| ||||| |||||||||||| |||
Db   15786 TTTTTTTTTTTAACATTTTTTTTTAATTTTTTTTATTTTTATTATATATATTTFTTAAT 15727

Qy   1689 ctggtggaacttcgataaatgagccatatctgagaatgtagctagctcaaaggaagtct 1748
           |||| |||| || || ||| |||||||||||| |||||| ||| ||||||||||
Db   68 tctatttatatatatgtatgtgttttaaaaatgtgtatatttatcaga---tttggt 2224

Qy   2285 gaattctagttccatcatatctattggaaagatataaaagaaagtataaggcaatca 2344
           ||| |||| |||||| |||| |||||||||||||||| ||| || |||| ||||
Db   15666 TTTTTTCTTTTCTTTTGTTTTTTTTATTTTTTTTATAAACATTTTTTTTTTTTT 15607

Qy   2225 tttgccttaaatattatccccaaattgcttcatcccattcgttcagtatatatatttt 2284
           ||||| |||||||| | ||||| |||| ||||| ||||| |||| |||||||
Db   15726 ATATTTTTTCTTTTTTTTTGTTTTTTATGATATATAATATATTTTAATTGTTTTTA 15667

Qy   68 tctatttatatatgtatgtgttttaaaaatgtgtatatttatcaga---tttggt 2224

Qy   1809 aatgtgagactcttaggaaactaactagatctccaattggagcacactaacaatagatgat 1868
           ||| |||||| ||||||| |||| |||| ||||||| ||| |||| |||| ||||
Db   300 ATTATAAATATTATTTAATTTAAACATTTTAATTTTATAAAATGTAATTTTTAAAACATTTTA 241

Qy   1749 gtcattaataaggtgatttcctaaattaaacacacattatgtaaggagagtgcaaaaatgt 1808
           ||||||| |||||| ||| | |||||| || || |||| |||| |||| |||| ||||
Db   360 ATTATAAATAATGTGTAATTTAAAACATTTTAATATATAAAATATGTAATTTAAACATTTTA 301

Qy   1929 atttattacttgaatttgaaactggactgtgttaggagttaaactataattctgtt 1988
           ||||||| |||| |||| || |||||| || | |||||||||| ||| ||||
Db   240 ATTAAAAAATTATTTTTAATTTTTTTAAAAAAATTTTTAATTTTTAAATTTATATTTTA 181

Qy   1869 accaaatactcttttgttttttaacacagctatccagtaaggctatcatgatgtgtgctaaa 1928
           ||||| ||| |||||||||||| |||| |||||||| || || ||||| || || |
Db   300 ATTATAAAATATTATTTAATTTAAACATTTTAATTTTATAAAATGTAATTTTTAAAACATTTTA 241

Qy   1629 tttaagaaaccaaagtccgtagtctaagataactaactgattaacatcttaaaaattatgtaata 1688
           ||||||||| ||||| || |||||||||| |||||||| |||||||| |||||| ||||||
Db   480 TGTAATTATTAAAAATACTTTATTAAAATATGTAATTTAAAATATGTAATTTTTA 421

Qy   2345 tatatttcattcaaaagatatttttagcaactgctatgtgtgcctttcgttgtt 2397
           ||||||| |||||| || |||||||| ||| ||| | ||| ||| || ||
Db   15546 AATTTTATTATATTTTTCTCCCTTTAAAAATTTTCCTTTTTTTTTTTTTT 15494

RESULT 3
US-08-487-826B-13
; Sequence 13, Application US/08487826B
; Patent No. 5993827
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhuan
; APPLICANT: Wellems, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 45

Best Local Similarity 46.4%; Score 56; DB 4; Length 665;
Matches 217; Conservative 0; Mismatches 250; Indels 1; Gaps 1;

Qy   1989 cctaaaagaaaatttatctgcaaagttctgagatattagctaatgaattagt 2048
           || ||| ||| ||||||||||| | ||| |||| |||| |||||| ||
Db   121 AATATAAAAATATTTATTAATTTTAATTTATATTTTTATATTTATAAAAATATTTTT 62

Qy   2049 gtttggggttactctcttgttctaagataagaatgtgaagaatatt 2096
           || |||||||| |||||| |||||||||| || ||| ||||||
Db   61 AATATAAAATATTTTTTAATTTAATTATATATAAATATAAAAACACAATT 14

; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 665 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: Rh 32
US-08-883-795A-36

Query Match 2.3%;
Best Local Similarity 46.4%; Pred. No. 0.00065;
Matches 217; Conservative 0; Mismatches 250; Indels 1; Gaps 1;

Best Local Similarity 46.9%; Pred. No. 0.00019;

CORRESPONDENCE ADDRESS:
ADDRESSEE: BERESKIN & PARR
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/883,795A
FILING DATE: 27-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gravelle, Micheline
REGISTRATION NUMBER: 40,261
REFERENCE/DOCKET NUMBER: 7841-062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398

CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
-08-487-826B-13

Query Match   2.2%; Score 54; DB 4; Length 19124;
Best Local Similarity   51.0%; Pred. No. 0.0043;
Matches 153; Conservative 0; Mismatches 145; Indels 2; Gaps 1;

```
Qy  1960 tgtgttaggggattaaactataattctgttcttaaaagaaaaatttatctgcaaatgtgcaa 2019
              |||||| | || |||||||   ||  | |||||  |       | ||
Db 17792 TGTGTAATTTTTTTTCATAATATTATATATATATATATATGTATTTTTTTAGTATAATAA 17851

Qy  2020 gttctgagatatataagctaatgaatta--gttgttgtgggttactttcttttgttctaagta 2077
              | |||  || |||| |  | || |  |||  ||||  |  ||  | | ||| ||
Db 17852 TTGTATCTATATTTGATTAATATATGTATATATATTATTTTTGTTTCTTTTTCTGTA 17911

Qy  2078 taagaatgtgaagagaatatttgaaaactcaatgaaataattctcagctgccaaatgttgca 2137
              | |  | |||| |||  |||   |||| ||   |  ||||||||||  |   |
Db 17912 CATTTTTCGTAATATATATATATATATATATATAATTCTCTTTTTTCTAATATATATA 17971

Qy  2138 ctcttttatatatcttcttccacttttgatctcatttatatatatggtatgtgtttttaaa 2197
              || |||| || ||||    ||||  | | ||||| || || || || ||||
Db 17972 TCCTTCTATTTCGATTTTTTCATTTTTTCCAGTATTAATTTATTTATTTATTTGTGAT 18031

Qy  2198 atatgtgtatatattcaagattggtttgtcttaaatattatcccaaattgcttcagt 2257
              | ||| | ||| ||   | |||  |  ||||  | |||| ||  |   |||
Db 18032 ATTTTATTATATATATTTTAAATGTGTATTTATTATATATATGTGTTTTTATTTTATT 18091
```

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTzgpt-F1s
US-08-232-463-14

Query Match   2.2%; Score 53.8; DB 1; Length 7218;
Best Local Similarity   4.0%; Pred. No. 0.0037;
Matches 16; Conservative 222; Mismatches 159; Indels 0; Gaps 0;

```
Qy  245 tattaagaaatataagcaattttggtcctgaagttacaaggcaacagactatccaactgt 304
             |||||  | ||  | ||||||| || |||:::::::::::::::::::::::::::::
Db 1462 TAGTTAAAGAGATAGAAGAAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1403

Qy  305 tgaggaaattcttaagaatcatgtaatgaaagatatcaaagagaggtggggatcaagaaa 364
            ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1402 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1343

Qy  365 atgttgatgataacaaccagctcttcagatttctgcaacttcgccacttaaaactctac 424
            ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1342 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1283

Qy  425 cacgaaggtatccaggaattgagaaaaaacaacatagagaaacttttccaaagataaagata 484
            :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1282 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1223

Qy  485 gcatttttaaattacgaaacttatctcgtagaactcctaaaaggcatggattacatttat 544
            :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1222 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1163

Qy  545 ctcaggaaatggcgagaaaaataaagccatgaaataatcaatgaagatcaagaaaatgcaa 604
            :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1162 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1103

Qy  605 ttgataatagagaactaagccaaggaagatgttgaaga 641
            ::::::::::::::::::::::::::::::::::::::|
Db 1102 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRA 1066
```

RESULT 5
US-08-613-861-2
Sequence 2, Application US/08613861
Patent No. 5843770
GENERAL INFORMATION:
APPLICANT: Ill, Charles R. et al.
TITLE OF INVENTION: Antisense Constructs Directed Against Viral Post-Transcript
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875

RESULT 4
US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:

COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/613,861
  FILING DATE: 13-APR-1994
  CLASSIFICATION: 514
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 08/111,111
  FILING DATE: 12-DEC-1909
ATTORNEY/AGENT INFORMATION:
  NAME: Attorney, Name Init
  REGISTRATION NUMBER: 000000
  REFERENCE/DOCKET NUMBER: oe
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (617)227-7400
  TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
  LENGTH: 4525 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: single
  TOPOLOGY: linear
  MOLECULE TYPE: cDNA
US-08-613-861-2

Query Match 2.1%; Score 52.4; DB 3; Length 4525;
Best Local Similarity 79.5%; Pred. No. 0.0065; Matches 62; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```
Qy    1 gcttggtaccgagctcggatccaactagtaacggccgccagtgtgctggaagggagaccca 60
        ||||||||||||||||||||||||||||||||||||||||||| |||||||||||| ||
Db 3090 GCTTGGTACCGAGCTCGGATCCAACTAGTAACGGCCGCCGACAGTGTGCTGGAATTCCTCAAGG 3149

Qy   61 ccgggcgctgacagacct 78
        ||| || ||||| || ||
Db 3150 TCGGTCGTTGACAATTGCT 3167
```

RESULT 6
US-08-458-023B-3
Sequence 3, Application US/08458023B
Patent No. 5667990
GENERAL INFORMATION:
  APPLICANT: Berka, Randy M.
  APPLICANT: Yoder, Wendy
  APPLICANT: Takagi, Shinobu
  APPLICANT: Boominathan, Karuppan C.
  TITLE OF INVENTION: ASPERGILLUS EXPRESSION SYSTEM
  NUMBER OF SEQUENCES: 16
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: NO. 5667990 No. 5667990disk of No. 5667990th America, Inc.
  STREET: 405 Lexington Avenue
  CITY: New York
  STATE: New York
  COUNTRY: USA
  ZIP: 10174-6201
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/458,023B
  FILING DATE: 01-JUN-1995
  CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
  NAME: Lowney Dr., Karen A.
  REGISTRATION NUMBER: 31,274
  REFERENCE/DOCKET NUMBER: 4086.010-US

TELECOMMUNICATION INFORMATION:
  TELEPHONE: 212-867-0123
  TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
  LENGTH: 1123 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: single
  TOPOLOGY: linear
  MOLECULE TYPE: DNA (genomic)
  FRAGMENT TYPE: internal
  ANTI-SENSE: NO
  HYPOTHETICAL: NO
  ORIGINAL SOURCE:
  ORGANISM: Humicola insolens
  INDIVIDUAL ISOLATE: DSM 6995
  FEATURE:
  NAME/KEY: CDS
  LOCATION: 126..806
US-08-458-023B-3

```
Qy    1 gcttggtaccgagctcggatccaactagtaacggccgccagtgtgctggaagggagagc 58
        ||||||||||||||||||||||||||||||||||||||||||| |||||||||| ||| ||
Db   28 GCTTGGTACCGAGCTCGGATCCAACTAGTAACGGCCGCCGACAGTGTGCTCTAAAGCCGCGC 85
```

Query Match 2.1%; Score 51.6; DB 1; Length 1123;
Best Local Similarity 93.1%; Pred. No. 0.0067; Matches 54; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

RESULT 7
US-07-867-106-2/c
Sequence 2, Application US/07867106
Patent No. 5389526
GENERAL INFORMATION:
  APPLICANT: Slade, Martin B
  APPLICANT: Chang, Andy C M
  APPLICANT: Williams, Keith L
  TITLE OF INVENTION: Improved Plasmid Vectors for Cellular
  TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium
  NUMBER OF SEQUENCES: 19
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris
  STREET: One Liberty Place 46th Floor
  CITY: Philadelphia
  STATE: PA
  COUNTRY: USA
  ZIP: 19103
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/07/867,106
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: AU PJ 7187
  FILING DATE: 19920625
  APPLICATION NUMBER: PCT/AU90/00530
  FILING DATE: 02-NOV-1989
ATTORNEY/AGENT INFORMATION:
  NAME: Feeney, Joanne Longo
  REGISTRATION NUMBER: 35,134
  REFERENCE/DOCKET NUMBER: RICE-0002
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 215-568-3100
  TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
  LENGTH: 5852 base pairs
  TYPE: NUCLEIC ACID
  STRANDEDNESS: single

```
;  TOPOLOGY:  linear
;  MOLECULE TYPE:  DNA (genomic)
;  ANTI-SENSE:  NO
;  FEATURE:
;    NAME/KEY:  CDS
;    LOCATION:  2378..5038
;  FEATURE:
;    NAME/KEY:  CDS
;    LOCATION:  2378..5038
;  US-07-867-106-2

Query Match           2.1%;  Score 50.6;  DB 1;  Length 5852;
Best Local Similarity 46.2%;  Pred. No. 0.017;
Matches  246;  Conservative  0;  Mismatches  279;  Indels  8;  Gaps  2;
```



```
Qy  1933  tatttacttgaatttgaaaactgagctgtgttaggattaaactatataattctgttctta 1992
          ||||||| || ||||||||||||| ||| |||| || | ||||| ||||||||| ||||||
Db  2153  TATATTATCCAATATTTATAAAATAAGGTATAATTAGATAGAGAATAAAAATTTTAAATTTA 2094

Qy  1993  aaagaaaatttatctgcaagttctgagatattagctaatcaatcgaattagttgtt 2052
          |||||                    ||  |||||  | | |   ||||||| |||||
Db  2093  TTAAAAAAATCAAAAAAATCAAAAACCAAAGTAATTAATATATTT-----ATAAATGAGGGTTTTT 2039

Qy  1993  aaagaaaatttatctgcaagttctgagatattagctaatcaatcgaattagttgtt 2052   (●)
          
Db  2093  

Qy  2053  ggggttactctcttgttcctaagtataagaaatgtgaagaagatacatttgaaaactcaatgaa 2112
          |||||||||||||| || ||||| | ||||| |||||||| ||||||||| |||||||||||
Db  2093  TTAAAAAAATCAAAAAAATCAAAAACCAAAGTAATTAATATATTT-----ATAAATGAGGGTTTTT 2039

Qy  2053  ggggttactctcttgttcctaagtataagaaatgtgaagaagatacatttgaaaactcaatgaa 2112
          
Db  2038  TTTTTTTTTTTTTTTTTTTTTTTCAAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAA 1979

Qy  2113  taattctccagcgtgccaaatgtgcacctcttttatatatttcttttccacttttgatctat 2172
          ||||| |||||            ||||||||||||||||||||||||||||||||||| ||||
Db  1978  TAGAAAAAAAGTTGGTTAAACTACATTAGTTTTTTATAGTTTTTTGCATATTTAAAAATAAC 1919

Qy  2173  ttatatatatgtgatgtgttttttaaaatatgtgtatattttatcagatttggttttttgcct 2232
          |||||||||                        ||||||||||||||||||||||||||||
Db  1978  TAGAAAAAAAGTTGGTTAAACTACATTAGTTTTTTATAGTTTTTTGCATATTTAAAAATAAC 1919

Qy  2173  ttatatatatgtgatgtgttttttaaaatatgtgtatattttatcagatttggttttttgcct 2232
          
Db  1918  TTTTAATTTAATTGATTTTTAATTATGAGATCTAATAAAAAAAAAAAAAATTTTTTAAAATT 1859

Qy  2233  aaatatattatcccaattgcttcagtcatccattgttcagtatatatatatttggaattcta 2292
          |||||| ||                  ||||||||||| ||||||||||||||||||||||||
Db  1858  TAAAAAAAAAAAGAAAAAAAAAAAGTAGAATTATTTAAAAATTAAAAATTATTATTATTCAA 1799

Qy  2293  gtttcatcataatctcattagaa---gatgggggataataaaagagaagtataaaggcaatcacatat 2349
          |||||||||| ||||||| ||||    |||| | |||| |||||||||||||||||| |||||
Db  1858  TAAAAAAAAAAAGAAAAAAAAAAAGTAGAATTATTTAAAAATTAAAAATTATTATTATTCAA 1799

Qy  2293  gtttcatcataatctcattagaa---gatgggggataataaaagagaagtataaaggcaatcacatat 2349
          
Db  1798  TCTTAATAAATTAAGTATAATATCGATAGGCAATTTATTTATTTTTATATCTACTAAAAAA 1739

Qy  2350  tcattcaaaagatatttattagcaactgctatgtgccttcgttgttccagatatgcag 2409
          ||||| ||||||||||||||||| |||   |||   ||||| ||||||||||||||||
Db  1738  CTAGGAAAAATGAAATGTCATCAAAATAGTATTTTAACATTTTTTTTTTTTTTTTTAAAA 1679

Qy  10    agacaatgataaaacatataatatctcttccaaaaaaaaaaaaaaaaaaaa 2462   (●)
          |||||||||||||||||||||||
Db  1678  AAAGTGTCATGACAAAAAAAAAAAAGTGTCATGACAAAAAAAAAAAAGTCGTCATGACAAA... 1626
```

```
RESULT   8
US-07-991-867B-8/c
;  Sequence 8, Application US/07991867B
;  Patent No. 5476781
;  GENERAL INFORMATION:
;    APPLICANT:  Moyer, Richard W.
;    APPLICANT:  Hall, Richard L.
;    APPLICANT:  Gruidl, Michael E.
;    TITLE OF INVENTION:  Entomopoxvirus Expression System
;    NUMBER OF SEQUENCES:  66
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  David R. Saliwanchik
;      STREET:  2421 N.W. 41st Street, Suite A-1
;      CITY:  Gainesville
;      STATE:  FL
;      COUNTRY:  USA
;      ZIP:  32606
;  COMPUTER READABLE FORM:
;    COMPUTER TYPE:  IBM PC compatible
;    MEDIUM TYPE:  Floppy disk
;    COMPUTER:  IBM PC compatible
```

Right column:

```
;  OPERATING SYSTEM:  PC-DOS/MS-DOS
;    SOFTWARE:  Patentin Release #1.0, Version #1.25
;  CURRENT APPLICATION DATA:
;    APPLICATION NUMBER:  US/07,991,867B
;    FILING DATE:  12-DEC-1992
;    CLASSIFICATION:  435
;  PRIOR APPLICATION DATA:
;    APPLICATION NUMBER:  WO 92/14818
;    FILING DATE:  12-FEB-1992
;  PRIOR APPLICATION DATA:
;    APPLICATION NUMBER:  US 07/827,685
;    FILING DATE:  30-JAN-1992
;  PRIOR APPLICATION DATA:
;    APPLICATION NUMBER:  US 07/657,584
;    FILING DATE:  19-FEB-1991
;  ATTORNEY/AGENT INFORMATION:
;    NAME:  Saliwanchik, David R.
;    REGISTRATION NUMBER:  31,794
;    REFERENCE/DOCKET NUMBER:  UF114.C3
;  TELECOMMUNICATION INFORMATION:
;    TELEPHONE:  904-375-8100
;    TELEFAX:  904-372-5800
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1511 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      ORGANISM:  Amsacta moorei entomopoxvirus
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  complement (18..218)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  CDS
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  complement (234..782)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  852..1511
;  US-07-991-867B-8

Query Match           2.0%;  Score 50.4;  DB 1;  Length 1511;
Best Local Similarity 46.5%;  Pred. No. 0.013;
Matches  199;  Conservative  0;  Mismatches  226;  Indels  3;  Gaps  1;

Qy  1873  aatacttttgtttttttaaacagctatccagtaaggctaccatccagtgatgtgtgctaaaatt 1932
          |||||||| |  | |||||||||||||||||||||||||| || | |||||||||||||||
Db  1243  ATTAAATTTTTTTAAAATTTACTAATATTATTATTAAAATTATAAGTGCATTTATTATTACATGAT 1184

Qy  1933  tatttacttgaatttgaaaactgagctgtgttaggattaaactatataattctgttctta 1992
          ||||||| ||  ||||||||||||| ||| |||||| ||||| |||| |||||||||||
Db  1243  ATTAAATTTTTTTAAAATTTACTAATATTATTATTAAAATTATAAGTGCATTTATTATTACATGAT 1184

Qy  1933  tatttacttgaatttgaaaactgagctgtgttaggattaaactatataattctgttctta 1992
          
Db  1183  TCACAATTAAAAAAATTCTATAGAATGTGGTAGTAATAAATTACTTATATAATTGCTATTTTG 1124

Qy  1993  aaagaaaatttatctgcaagttctgagatattagctaatcaatcgaattagttgtt 2052
          |||||| |                  ||  ||||| | | |   ||||||| |||||
Db  1123  TTATTAAGAATATATCTAAATATGTTAAAATGTTTTTAATTGTTTATATAAATTAAATTAATA 1064

Qy  2053  ggggttactctct---ttgttcctaagtataagaatgtgaagaagatatttgaaaactcaatg 2109
          ||||||||| |     ||||||| ||||| | ||| | ||||||||| ||||| ||||||||
Db  1063  ATATTTAAATTGAAATAAATTAAAACTTTTAATAATATTTCTGGAATATTTATTTAAATATTA 1004

Qy  2110  aaataaatttcagtgccaaatgttccactcttttatatattcttttccacttttgatc 2169
          ||||| | |||||| | ||||||| |||| ||||||||| ||| |  | ||||| |||
Db  1003  TTATCATAATAATATTATATGCAATTCTTCTAAAATTAACTAAATTTTTTAATAATATTAATA 944

Qy  2170  tatttatatatgtgtgttttttaaaatatgtgtatattttatcagatttggttttttgc 2229
          |||| |||||||| ||  ||||| ||||| | |||  ||| ||||||||| ||||||
Db  943   TTAATAACATTATCTCGTTTATTTATTATATTTATTTATTTTTTAATTATATAATAATATTT 884

Qy  2230  cttaaatatatcccaattgcttcagtcattcattgttcagtatatatatatttgaatt 2289
          ||||| ||      |||||||||| ||||||| ||||| |||||||||||||  | |||
Db  943   TTAATAACATTATCTCGTTTATTTATTATATTTATTTATTTTTTAATTATATAATAATATTT 884
```

RESULT 9
US-08-107-755A-8/c
Sequence 8, Application US/08107755A
Patent No. 5721352
GENERAL INFORMATION:
APPLICANT: Moyer, Richard W.
APPLICANT: Hall, Richard L.
APPLICANT: Gruidl, Michael E.
TITLE OF INVENTION: No. 5721352el Entomopoxvirus Expression System
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: U.S.A.
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/107,755A
FILING DATE: 19-AUG-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,658
FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/657,584
FILING DATE: 19-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF114.C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1511 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Amsacta moorei entemopoxvirus
FEATURE:
NAME/KEY: CDS
LOCATION: complement (18..218)
FEATURE:
NAME/KEY: CDS
LOCATION: complement (234..782)
FEATURE:
NAME/KEY: CDS
LOCATION: 852..1511

Query Match   2.0%; Score 50.4; DB 2; Length 1511;
Best Local Similarity   46.5%; Pred. No. 0.013;
Matches 199; Conservative   0; Mismatches 226; Indels   3; Gaps   1;

Qy   1873 aatactttttgttttaaccagctcatccagtgtaaggctatccatcatgatgtgtgctaaattt 1932
          | || |||||| || |||| |||| || || |||| ||| |||| || | |||| || |

Db   883 ATTAAAATTATATCAGAATTTAGTAAAATCCATTTGATAAATTTATTTTTTTCATTGA 824
          ||||| |||
Qy   2290 ctagtttt 2297
          |||||
Db   823 TTAATTT 816

Db   1243 ATTAATTTTTTAAATTTACTAAATTATTAATAAAATTATAGTCATTTATTATTACATGAT 1184
Qy   1933 tatttacttgaatttgaaaactgagctgtgttaggggattaaaactataattctgtttctta 1992
Db   1183 TCACAATTTAAAATTCTATAGAAGTGTGGTAGTATAAATAATTACTTATATATGCATTTG 1124
Qy   1993 aaagaaaatttactctgcaaatgtgcaagttctgagatattagctaatgaattagttgttt 2052
Db   1123 TTATAAGATATATCTAAATATGTTATATATTTTTAATTTGTTTATAGAAACATTAAAATAAA 1064
Qy   2053 ggggttacttct---ttgttctaagtataagaaagtgaagaaatttgaaaactcaatg 2109
Db   1063 ATATTTAAATTTGAAATATATAAAACTTTAAATATTTTCTCGGAATATATATTTAAAATATTA 1004
Qy   2110 aaataattcctcagctgccaaatgttgcactcctctttttatatattcttttccacttttgatc 2169
Db   1003 TTATCATAAATATATATATGGAATCTTCTTCTAAAATTAAACTAATTTTTTTTAAATATATATA 944
Qy   2170 tatttatatatatgtatgtgtttttaaaatatgtgtatatattttatccagatttgggttttgc 2229
Db   943 TTAATAACAATATCTCGTTATTATTTATTATTTTTTAAAATATATAATTATTTAAAATATTT 884
Qy   2230 cttaaatatatatcccaaatgcttcagtcattgttcagtatatatatatttgaatt 2289
Db   883 ATTAAAATTATATCAGAATTTAGTAAAATCCATTTGATAAATTTATTTTTTTTCATTGA 824
Qy   2290 ctagtttt 2297
Db   823 TTAATTT 816

RESULT 10
US-08-544-332-8/c
Sequence 8, Application US/08544332
Patent No. 5935777
GENERAL INFORMATION:
APPLICANT: Moyer, Richard W.
APPLICANT: Hall, Richard L.
APPLICANT: Gruidl, Michael E.
TITLE OF INVENTION: No. 5935777el Entomopoxvirus Expression System
NUMBER OF SEQUENCES: 77
CORRESPONDENCE ADDRESS:
ADDRESSEE: Gerard H. Bencen
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/544,332
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/991,867
FILING DATE: 07-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/107,755
FILING DATE: 19-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 92/14818
FILING DATE: 12-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,685
FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/657,584
FILING DATE: 19-FEB-1991
ATTORNEY/AGENT INFORMATION:

; NAME: Bencen, Gerard H.
; REGISTRATION NUMBER: 35,746
; REFERENCE/DOCKET NUMBER: UF114.C4
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 904-372-5800
;   TELEFAX: 904-375-8100
; INFORMATION FOR SEQ ID NO: 8:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1511 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: unknown
;   MOLECULE TYPE: DNA (genomic)
;   ORIGINAL SOURCE:
;   ORGANISM: Amsacta moorei entomopoxvirus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: complement (18..218)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: complement (234..782)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 852..1511
; US-08-544-332-8

Query Match   2.0%; Score 50.4; DB 4; Length 1511;
Best Local Similarity 46.5%; Pred. No. 0.013;
Matches 199; Conservative 0; Mismatches 226; Indels 3; Gaps 1;

Qy 1873 aatactttttgtttttaacacagctaccccagtaaaggctcatcatgatgtgtgctaaattt 1932
              | ||||||||||||| ||||| || ||||||| |||| ||||||| || ||||||
Db 1243 ATTAATTTTTTATTACTAAAATTATTAATAAAATTATATGTCATTTATTATTACATGAT 1184

Qy 1933 tatttacttgaatttgaaaactgagctgttgttaggagattaaactatataattctgttctta 1992
             | |||||||| ||||||||||| |||||||| || ||| ||||| ||||| ||||||
Db 1183 TCACAATTTAAAAATTCTATAGAATGTGGTAGTATATAAATTACTATTATTATTGCTATTTTTG 1124

Qy 1993 aaagaaaatttatctgcaaatgtgcaagttctggagatattagctaatgaattagttgtt 2052
             ||| | ||||||||||||| ||||| ||| |||||| | ||||||| ||||| ||||
Db 1123 TTATAAGATATAAATAAATAAATGTATAATATATTTTTTAATTTTGTTATAAAAATTAAAATTAATA 1064

Qy 2053 gggttacttct---ttgtttctaagtatataagaaatggaagaatatttgaaaactcaatg 2109
              ||| |||| | ||||| ||||| ||||| |||||||||| ||| |||||||||||
Db 1063 ATATTTAAATTTGAAATATAATATAAAACTTTTAATATTTCTGGAATAATTATTTAAAATATTA 1004

Qy 2110 aaataattctcagctgccaaaatgttgcactctttttatatattcttttccacttttgatc 2169
             |||| | ||||| ||| | || ||| || |||||| || |||||| || ||||||
Db 1003 TTATCATATATATATTATATGCAATTCTCTAAATAACTAATTTTTTTTAATATTAATTA 944

Qy 2170 tatttatatatatgtatgtgttttaaaatatgtgtatcatttatcagattggttttgc 2229
             || | |||||||| || | ||||||||||||| |||||||| ||||||| ||||||
Db 943 TTAATAACATTAACTCTGTTTATATTATTATTTTTTATTATTATTAATATATTTAAAATATTT 884

Qy 2230 cttaaaatatattaccccaattgcttcagtcattcattgttcagtatatatatttgaatt 2289
             ||||| | ||||||| | | |||||||| | |||| ||| || |||||||| ||| ||
Db 883 ATTAAAATTATATCAGAATTTAGTAAGTAAATCCATTTTGATAATTTTATTTTTTTCATTGA 824

Qy 2290 ctagtgttt 2297
             || |||||
Db 823 TTAATTTT 816

RESULT 11
US-08-832-883-68
; Sequence 68, Application US/08832883
; Patent No. 5807651
; GENERAL INFORMATION:
;   APPLICANT: Giordano, Antonio
;   APPLICANT: Baldi, Alphonso
;   TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS AND PROGNOSIS
;   TITLE OF INVENTION: OF CANCER

NUMBER OF SEQUENCES: 115
CORRESPONDENCE ADDRESS:
;   ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.
;   STREET: Suite 1800 Two Penn Center Plaza
;   CITY: Philadelphia
;   STATE: PA
;   COUNTRY: USA
;   ZIP: 19102
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/832,883
;   FILING DATE:
;   CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
;   NAME: Monaco, Daniel A
;   REGISTRATION NUMBER: 30,480
;   REFERENCE/DOCKET NUMBER: 8321-13 US1
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (215) 568-8383
;   TELEFAX: (215) 568-5549
; INFORMATION FOR SEQ ID NO: 68:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 9837 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
; US-08-832-883-68

Query Match   2.0%; Score 50.2; DB 2; Length 9837;
Best Local Similarity 81.7%; Pred. No. 0.024;
Matches 58; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1 gcttggtaccgagctccggaatcccacactaagtaacggccgccagtgtgctggaaaggggacgcca 60
          |||||||||||||||||||| ||||||||||| ||||||||||||||||| | ||||||||||||
Db 7173 GCTTGGTACCGAGCTCGGGATCCACTACTAGTAACGGCCGCCAGTGTGCTGGAATTCGGCTTAA 7232

Qy 61 ccgggcgctga 71
          || || ||||
Db 7233 AGGTAGGCCGGA 7243

RESULT 12
US-08-832-877-68
; Sequence 68, Application US/08832877
; Patent No. 5840506
; GENERAL INFORMATION:
;   APPLICANT: Giordano, Antonio
;   TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS AND PROGNOSIS OF
;   TITLE OF INVENTION: CANCER
;   NUMBER OF SEQUENCES: 116
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.
;   STREET: Suite 1800 Two Penn Center Plaza
;   CITY: Philadelphia
;   STATE: PA
;   COUNTRY: USA
;   ZIP: 19102
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/832,877
;   FILING DATE:
;   CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:

NAME: Monaco, Daniel A
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 8321-13 US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 9837 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

US-08-832-877-68

Qy 61 ccgggcgctga 71
Db 7233 AGGTAGGCGGA 7243

Qy 1 gcttggtaccgagctccggatccacactagtaacggrccgccagtgtgctggaaggggacgccca 60
       |||||||||||||||| ||||||||||||||||||||||||||||||||||||| || ||||
Db 7173 GCTTGGTACCGAGCTCGGATCCACACTAGTAACGGCCCCGCCCAGTGTGCTGCGAATTCGGCTTAA 7232

RESULT 13
US-08-040-753-1
Sequence 1, Application US/08040753
Patent No. 5464745
GENERAL INFORMATION:
APPLICANT: Mierendorf, Robert
APPLICANT: Garber, Richard
APPLICANT: No. 5464745y, Robert
APPLICANT: Hammer, Beth
TITLE OF INVENTION: Protein Ligand Binding
TITLE OF INVENTION: Region Mapping System
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles and Brady
STREET: 1 South Pinckney St., Box 2113
CITY: Madison
STATE: WI
COUNTRY: USA
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/040,753
FILING DATE: 19930331
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 70-399-9001-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3984 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO

ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: pTOPE-1b(+)
FEATURE:
NAME/KEY: promoter
LOCATION: 67..85
OTHER INFORMATION: /function= "T7 Promoter"
FEATURE:
NAME/KEY: RBS
LOCATION: 134..139
FEATURE:
NAME/KEY: CDS
LOCATION: 148..1092
OTHER INFORMATION: /product= "5' end of T7 gene 10
OTHER INFORMATION: fusion protein"/note= "This coding region is interrupted
OTHER INFORMATION: during cloning by insertion of putative epitope encoding
OTHER INFORMATION: DNA."

US-08-040-753-1

Qy 1 gcttggtaccgagctccggatccacactagtaacggrccgccagtgtgctggaa 50
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 938 GCTTGGTACCGAGCTCGGATCCACACTAGTAACGGCCCCGCCCAGTGTGCTGGAA 987

Query Match 2.0%; Score 50; DB 1; Length 3984;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 14
US-08-252-493C-4
Sequence 4, Application US/08252493C
Patent No. 5891645
GENERAL INFORMATION:
APPLICANT: Rollins, Scott
APPLICANT: Rother, Russell P.
APPLICANT: Evans, Mark J.
APPLICANT: Matis, Louis A.
TITLE OF INVENTION: PORCINE E-SELECTIN
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seth A. Fidel
STREET: 25 Science Park, Box 15
CITY: New Haven
STATE: Connecticut
COUNTRY: USA
ZIP: 06511
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 750 Kb storage
COMPUTER: PC compatible
OPERATING SYSTEM: DOS 6.2
SOFTWARE: WordPerfect 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/252,493C
FILING DATE: June 1, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Fidel, Seth A.
REGISTRATION NUMBER: 38,449
REFERENCE/DOCKET NUMBER: ALX-138
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 776-1790
TELEFAX: (203) 772-3655
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 4059 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Circular
MOLECULE TYPE: Other nucleic acid

Query Match 2.0%; Score 50.2; DB 3; Length 9837;
Best Local Similarity 81.7%; Pred. No. 0.024;
Matches 58; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```
;
;       DESCRIPTION:  Apex-1 Eukaryotic
;       DESCRIPTION:  Expression Vector
US-08-252-493C-4

Qy      1 gcttggtaccgagctcggatccactagtaaccggccgccgccagtgtgctggaa 50
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1027 GCTTGGTACCGAGCTCGGATCCACTAGTAACCGGCCGCCGCCAGTGTGCTGGAA 1076

        Query Match          2.0%;   Score 50;   DB 3;   Length 4059;
        Best Local Similarity  100.0%;   Pred. No. 0.021;
        Matches 50;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1 gcttggtaccgagctcggatccactagtaaccggccgccgccagtgtgctggaa 50
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1027 GCTTGGTACCGAGCTCGGATCCACTAGTAACCGGCCGCCGCCAGTGTGCTGGAA 1076
```

```
RESULT 15
PCT-US95-07554-4
;Sequence 4, Application PC/TUS9507554.
;GENERAL INFORMATION:
;       APPLICANT:  Sandrin, Mauro S.
;       APPLICANT:  Fodor, William L.
;       APPLICANT:  Rother, Russell P.
;       APPLICANT:  Squinto, Stephen P.
;       APPLICANT:  McKenzie, Ian F. C.
;       TITLE OF INVENTION:  Methods for Reducing
;       TITLE OF INVENTION:  Hyperacute Rejection of Xenografts
;       NUMBER OF SEQUENCES:  5
;       CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Maurice M. Klee
;       STREET:  1951 Burr Street
;       CITY:  Fairfield
;       STATE:  Connecticut
;       COUNTRY:  USA
;       ZIP:  06430
;       COMPUTER READABLE FORM:
;       MEDIUM TYPE:  3.5 inch, 750 Kb storage
;       COMPUTER:  Dell 486/50
;       OPERATING SYSTEM:  DOS 6.2
;       SOFTWARE:  WordPerfect 6.0
;       CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US95/07554
;       FILING DATE:
;       CLASSIFICATION:
;       PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/260,201
;       FILING DATE:  June 15, 1994
;       CLASSIFICATION:
;       APPLICATION NUMBER:  08/278,282
;       FILING DATE:  July 21, 1994
;       CLASSIFICATION:
;       ATTORNEY/AGENT INFORMATION:
;       NAME:  Klee, Maurice M.
;       REGISTRATION NUMBER:  30,399
;       REFERENCE/DOCKET NUMBER:  ALX-144.1PCT
;       TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (203) 255-1400
;       TELEFAX:  (203) 254-1101
;       INFORMATION FOR SEQ ID NO: 4:
;       SEQUENCE CHARACTERISTICS:
;       LENGTH:  4059 base pairs
;       TYPE:  Nucleic Acid
;       STRANDEDNESS:  Double
;       TOPOLOGY:  Circular
;       MOLECULE TYPE:  Other nucleic acid
;       DESCRIPTION:  Apex-1 Eukaryotic
;       DESCRIPTION:
;       ANTI-SENSE:  NO
PCT-US95-07554-4

        Query Match          2.0%;   Score 50;   DB 6;   Length 4059;
        Best Local Similarity  100.0%;   Pred. No. 0.021;
        Matches 50;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

```
Db   1027 GCTTGGTACCGAGCTCGGATCCACTAGTAACCGGCCGCCGCCAGTGTGCTGGAA 1076

        Search completed: March 24, 2000, 13:23:00
        Job time: 8171 sec
```

```
LOCUS       AA213817        456 bp    mRNA          EST       13-AUG-1997
DEFINITION  zr91d11.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:683061 3',
            mRNA sequence.
ACC    ION   AA213817
VE         AA213817.1  GI:1812444
KEY   DS   EST.
SOURCE      human.
```

```
ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
           Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1  (bases 1 to 456)
  AUTHORS  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE    National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
           Tumor Gene Index
  JOURNAL  Unpublished (1997)
COMMENT    On Apr 14, 1993 this sequence version replaced gi:315340.
           Contact: Robert Strausberg, Ph.D.
           Tel: (301) 496-1550
           Email: Robert_Strausberg@nih.gov
           This clone is available royalty-free through LLNL ; contact the
           IMAGE Consortium (info@image.llnl.gov) for further information.
           Insert Length: 1729   Std Error: 0.00
           Seq primer: -41m13 fwd. ET from Amersham
           High quality sequence stop: 317.
FEATURES          Location/Qualifiers
     source       1. .456
                  /organism="Homo sapiens"
                  /db_xref="GDB:5586381"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:683061"
                  /clone_lib="NCI_CGAP_GCB1"
                  /tissue_type="germinal center B cell"
                  /lab_host="DH10B"
                  /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                  polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                  was prepared from human tonsillar cells enriched for
                  germinal center B cells by flow sorting (CD20+, IgD-),
                  provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
                  (NCI) and Dr. Gerald Marti (CBER).  cDNA synthesis was
                  primed with a Not I - oligo(dT) primer
                  [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCCTCATTTTTTTTTTTTTTTTTT-
                  3']. Double-stranded cDNA was ligated to Eco RI adaptors
                  (Pharmacia), digested with Not I and cloned into the Not I
                  and Eco RI sites of the modified pT7T3 vector. Library
                  went through one round of normalization, and was
                  constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT      141 a     88 c      86 g     139 t      2 others
ORIGIN
```

```
Query Match              14.5%;  Score 356.8;  DB 30;  Length 456;
Best Local Similarity    95.2%;  Pred. No. 2.7e-57;
Matches  432; Conservative    0; Mismatches   14; Indels    8; Gaps    6;

Qy   1054 cccccaccaaatcgtagaaagcttcaacttttaatgcgtatgattt-cccgaatgagtca 1112
          |||   | || |||||||||   |  ||||||||||||||||||||| |||||||||||||
Db    452 CCCAACCAAANTCGTAGAAGCTTCAAACTTTTAATGCGTATGATTTCCCCGAATGAGTCA 393

Qy   1113 aaat--gttgatatgcccaaactt-catgatgcaatgggtacgaggtcac-tgatgatac 1168
          |||   || ||||||||||||||| ||||||||||||||||||||||||| ||||||| |
Db    392 AAAATGGTGGATATGCCCAAACTTCCATGATGCAATGGGTACGAGGTCACTTGATGATNC 333

Qy   1169 atacctttctcgatgtgtgttatgctgtgctgaagaagtggatcttgatgagcttcttg 1228
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    332 ATACCTTTTCTCGATGTGTGTTATGCTGTGCTGAAGAAGTGGATCTTGATGAGCTTCTTG 273

Qy   1229 ctggaagattagtttctttcttaatggatcatcatcaggaaattcttcaagtaccctctt 1288
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    272 CTGGAAGATTAGTTTCTTTCTTAATGGATCATCATCAGGAAATTCTTCAAGTACCCTCTT 213

Qy   1289 acttacagactgcagtggaaaaacatcttgactacttaaaaaagggacatattgaaaatc 1348
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    212 ACTTACAGACTGCAGTGGAAAAACATCTTGACTACTTAAAAAAGGGACATATTGAAAATC 153

Qy   1349 ctggagatggactatttgctcctttgccaacttactcatactgtaagcagattagtgctc 1408
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    152 CTGGAGATGGACTATTTGCTCCTTTGCCAACTTACTCATACTGTAAGCAGATTAGTGCTC 93

Qy   1409 aggagtttgatgagcaaaaagtttctacctctcaa-gctgcaattgcagaacttttagaa 1467
          |||||||||||||||| |||||||||||||||||| ||||||||||||||||||||||||
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2000, 21:51:58 ; Search time 3853.81 Seconds
                                      (without alignments)
                                      2413.056 Million cell updates/sec

Title: US-09-183-789-43
Perfect score: 2463
Sequence: 1 gcttggtaccgagctcggat.........aaaaaaaaaaaaaaaaaaaa 2463

Scoring table: IDENTITY_NUC

Searched: 4538634 seqs, 1887831982 residues

base :   EST:*
Word size :   0

Number of hits that pass the threshold   :   9077268

1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*

50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_gss1:*
80: gb_gss2:*
81: gb_gss3:*
82: gb_gss4:*
83: gb_gss5:*
84: em_gss1:*
85: em_gss2:*
86: em_gss3:*
87: gb_gss6:*
88: gb_gss7:*
89: gb_gss8:*
90: gb_gss9:*
91: gb_gss10:*
92: em_gss4:*
93: em_gss5:*
94: em_gss6:*
95: em_gss7:*
96: em_gss8:*
97: em_gss9:*
98: em_gss10:*
99: em_gss11:*
100: gb_gss11:*
101: em_gss12:*
102: gb_gss12:*
103: gb_gss13:*
104: gb_gss14:*
105: gb_gss15:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Pred. No.	Result No.	Score	% Query Match	Match Length	DB	ID	Description
C	1	519.2	21.1	559	41	AI052728	AI052728 ozz27b05.x
C	2	504.2	20.5	639	100	AQ280053	AQ280053 CITBI-E1-
C	3	501.2	20.3	517	38	AA807217	AA807217 oc35g09.s
C	4	499.8	20.3	503	61	AI810084	AI810084 wf65c09.x
C	5	362.6	14.7	692	35	AA545060	AA545060 vj91a10.s
C	6	356.8	14.5	456	30	AA213817	AA213817 zr91d11.s
C	7	326	13.2	328	38	AA756999	AA756999 ah54b09.s

RESULT 1

LOCUS   AI052728   559 bp   mRNA   EST   01-OCT-1998
DEFINITION 0227b05.x1 Soares_total_fetus_N2bHF8_9w Homo sapiens cDNA clone
            IMAGE:1676529 3', mRNA sequence.
ACCESSION   AI052728
VERSION   AI052728.1 GI:3308719
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 559)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   On Jan 19, 1998 this sequence version replaced gi:2152886.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 2396   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 492.
FEATURES   Location/Qualifiers
     source   1..559
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"

No.					Accession	Clone
8	325	13.2	482	39	AA825936	od59h06.s
9	312.4	12.7	698	102	AA376588	RPCI-11-1
10	271	11.0	387	33	AA420348	vd01g02.s
11	269.4	10.9	361	39	AA907054	om1g08.s
12	258.2	10.5	378	100	AQ261251	CITBI-E1-
13	206.2	8.4	496	87	AQ791096	HS_2189_B
14	185	7.5	501	102	AA240712	mv30b11.r
15	174.2	7.1	564	30	AA415970	RPCI-11-2
16	173.4	7.0	654	70	AW153497	f122d05.y
17	166.4	6.8	405	36	AA603375	zu71e10.5
18	161.8	6.6	656	104	AQ534381	RPCI-11-3
19	138	5.6	612	69	AW133819	f112c01.y
20	131	5.3	620	70	AW154095	f122d05.x
21	130.2	5.3	185	30	AA213896	zr91d11.r
22	110.8	4.5	327	64	AL118219	RPCI-11
23	106.4	4.3	259	30	R78674	yi74d03.r1
24	100.8	4.1	259	22	AA254513	val7c12.r
25	96.8	3.9	1101	79	CNS00EVL	Drosophil
26	94.2	3.8	1101	80	CNS0145U	Drosophil
27	87.2	3.5	1101	79	CNS0039G	Drosophil
28	86.2	3.5	258	47	AI508423	mv30b11.y
29	84	3.4	1101	79	CNS0039G	Drosophil
30	80.8	3.3	1101	79	AI063921	Drosophil
31	80	3.2	1101	79	CNS003BD	Drosophil
32	80	3.2	1101	79	CNS03DO1	Drosophil
33	78.8	3.2	1101	79	CNS00LT2	Drosophil
34	78	3.2	320	22	CNS0021J	Drosophil
35	76.6	3.1	1101	79	H01708	Drosophil
36	76.6	3.1	876	79	CNS00EVL	Drosophil
37	74.8	3.0	876	79	CNS009G1	Drosophil
38	74.4	3.0	341	79	CNS00O6	Drosophil
39	74	3.0	843	79	CNS00CS1	Drosophil
40	74	3.0	928	79	CNS00DKY	Drosophil
41	73.6	3.0	1001	80	CNS0155H	Drosophil
42	73.4	3.0	1101	79	CNS0021J	Drosophil
43	72.6	2.9	1101	79	CNS00LT2	Drosophil
44	72.2	2.9	1043	80	CNS0145P	Drosophil
45	72	2.9	1225	80	CNS0161D	Drosophil
	71.6	2.9	843	79	CNS00CS1	Drosophil

BASE COUNT   189 a   113 c   123 g   133 t   1 others
ORIGIN

ALIGNMENTS

Query Match   21.1%;   Score 519.2;   DB 41;   Length 559;
Best Local Similarity   99.2%;   Pred. No. 1.1e-87;
Matches 521; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy   51 agggacgcacacgggcgctgacaacagccctatggaaagtcaggggtgtgtcctcccgggccctta 110
         |||||||||||||||||||||||||||||||| |||||||| ||||||||||| |||| ||||||||||
Db   35 ATGGAGCGCCACCGGGCGCTGACAACAGACCTATGGAGAGTCAGGGTGTGCCCTCCGGGCCCTTA 94

Qy  111 tcgggcccaccaagctgtgaatgaagttaccacatctttcgagcaggaaagtgcctctaag 170
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   95 TCGGGCCCACCAAGCTGTGAATGAAGTTACCACATCTTTCGAGCAGGAAAGTGCCTCTAAG 154

Qy  171 aaaacacgacacaacactttaaaaaatatggcaattgtgtttcacagcagaggagacagtgga 230
         |||||||| |||||||||||||||| |||||||||||||||||||||||| ||||| ||||||||
Db  155 AAAACACAGACACAACACTTTAAAAAAATATGGCAATTGTGTTTCACAGCGGAGGAAGCAGTGGA 214

Qy  231 ttggctttatgacctattaagaaaataaatagcaatttggtccctgaagttacaaggcaaca 290
         |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db  215 TTGGCTTTATGACCTATTAAGAAAATAAATAGCAATTTGGTCCTGAAGTTACAAGGCAACA 274

Qy  291 gactatccaactgttgagaaaatttcttaagaaatacatgtaattcgaaagatatcaaaggaag 350
         |||||||||||||||||||||||||||||||||| ||||||| |||||||||||||||||| |||
Db  275 GACTATCCAACTGTTGAGAAAATTTCTTAAGAAATCAATGTAAATTCGAAAGATATCAAAGGGAG 334

Qy  351 gtgggggatcagaaaatgttgatgataacacccagcagctcttcagattcctcgcaacttcgcc 410
         ||||| |||||||||||||||||||||||||||||||||||||||| ||| |||||||||||||
Db  335 GTGGGGATCAGAAAATGTTGATGATAACACCCAGCAGCTCTTCAGAATTTCTGCAACTTCGCC 394

Qy  411 acttaaaaactctaccacggaggtatccagaattgagaaaaacaacatagagagaacttttc 470
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  395 ACTTAAAAACTCTACCACGGAGGTATCCAGAATTGAGAAAAACAACATAGAGAGAACTTTTC 454

Qy  471 caaagataaagatagcatttctttaaattacgaaaacttatcctgtagaaacttcctaaaaggca 530
         ||||||||||||||| |||| ||||||||||||||||||||| |||||||||||||||||||||
Db  455 CAAAGATAAAGATAGCCATTTTTAAATTACGAAAACTTATCTCGTAGAAACTTCCTAAAAGGCA 514

Qy  531 tggattacatttatctcaggaaaatggcgagaaaataaagcatga 575
         |||||||||||||||||| |||||||||| ||||||||||||||||
Db  515 TGGATTACATTTATCTCANGAAAAATGGCCAGAAAATAAAGCATGA 559
```

RESULT 2

LOCUS   AQ280053   639 bp   DNA   GSS   22-NOV-1998
DEFINITION CITBI-E1-251iN5.TR CITBI-E1 Homo sapiens genomic clone 251iN5,
            genomic survey sequence.
ACCESSION   AQ280053
VERSION   AQ280053.1 GI:3905872
KEYWORDS   GSS.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 639)
AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,

Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
Venter,J.C.
TITLE   Use of a random human BAC End Sequence Database for Sequence-Ready
           Map Building
JOURNAL   Unpublished (1998)
COMMENT   Other GSSs: CITB1-E1-251IN5.TF
           Contact: Mark Adams
           The Institute for Genomic Research
           9712 Medical Center Dr., Rockville, MD 20850, USA
           Tel: 301 838 0200
           Fax: 301 838 0208
           Email: mdadams@tigr.org
           Clones are available from Research Genetics (info@resgen.com), BAC
           end search page:
           http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
           Seq primer: M13 Reverse
           Class: BAC ends.
FEATURES   Location/Qualifiers
     source   1..639
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="251IN5"
                     /clone_lib="CITBI-E1"
                     /sex="male"
                     /cell_type="sperm"
                     /note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI;
                     Caltech Human BAC Library D"
BASE COUNT   186 a   102 c   107 g   244 t
ORIGIN

Query Match   20.5%;   Score 504.2; DB 100; Length 639;
Best Local Similarity   95.7%;   Pred. No. 7.1e-85;
Matches 518; Conservative   0; Mismatches   23; Indels   0; Gaps   0;

Qy 448 aaaacaacacatagagagaacttttcccaaagataaagatagcattttaaattacgaaactta 507
        ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db 639 AAAAACAACACATAGAGAAGAACTTTCCAAAGATAAAGATAGTATTTTAAATGATGAAACTTA 580

Qy 508 tctcgtagaactcctaaaagcatggatttacattttatctcaggaaaatggcgagaaaata 567
        |||| |||||||||||||| |||||||||||||||||||||||||||| |||||||||||
Db 579 TCTCATAGAACTCCTAAAAAGCATGGATTTACATTTATCTCAGGAAAATTCCGAGAAAATA 520

Qy 568 aagcatgaaataatcaattcaagataaaaatgcaattgataataataagaagaactaagccag 627
        ||||||||||||||||||| ||||||||||||||| |||||||||| ||||||||||||||
Db 519 AAGCATGAAATAATCAATGAAGAACAAGAAAAATGCAGTTGATAATAGAAAAACTAAGCCAG 460

Qy 628 gaagatgttggaagaagtttggagagatatgttattctgatctacctgcaaaacattttaggt 687
        |||||||||| ||||| ||||||| |||||||||||||||||||||||||| ||||||||||
Db 459 GAAGATGTTGAAGAAATTTGGAGATATGTTATTCTGATCTACCTGCAAACCATTTTAGGT 400

Qy 688 gtgccaatccctagaagaagtcaataaaatccaaaacaaagtaattccccaatatatataatgttac 747
        ||||| |||||||||||||||||| ||||||||| ||||||||||| |||||||| |||||||||
Db 399 GTGCCATCCCTAGAAGAAGTCATAAAATCCAAACAAGTAAATTCCCAATATATAATGTAC 340

Qy 748 aacatgtgccaataacaaagtaaacgtggagtagttatactacaaacaaaataacaagatgacctc 807
        ||||| |||||||||||||||||| ||||||||||| ||||||||| ||||||| ||||||||||
Db 339 AACATGGCCAATAACAAGTAAACATGGAGTAGTTACACTACAAAAACAAAATCAGATGACCTC 280

Qy 808 cctcactggttattatctggccatgaaagtgcctagacaaaattggccaagaagacaatgatatg 867
        ||||||||| ||||||||| |||||||||||||||| |||||||||||||||||| ||||| |||
Db 279 CCTCACTGGGTATTATCTGCCATGAAAGTGCCTAGCAAAATTGGCCAAGAAGTGATGAATACG 220

Qy 868 aataatccaacttatgttgttgaaccagagttcaagaaaatcccagattatttt 927
        ||||||||||||||||||| | ||| ||| | ||| |||| ||| ||||||||
Db 219 AATAATCCAACTTATGTTGGATTTGAACAAGATATATTCAGCACACAATCACAGATTATTTT 160

Qy 928 ctagatctccctgaacctcacttacttctttgtaattacgaattatttgtaaacattttg 987
        |||||||||| |||||||| | ||| |||| |||||||| |||||||||| ||| |||||
Db 159 CTAGATCTCCTTGAACCTCTCCTTACTTTACTTTGTAATTATTATGAATTATTTGTGAACANTTTG 100

Qy 988 g 988
        | |
Db g 988

Db 99 G 99
      |
   99

RESULT 3
LOCUS   AA807217   517 bp   mRNA   EST   07-APR-1998
DEFINITION oc35g09.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1351744 3',
           mRNA sequence.
ACCESSION AA807217
VERSION   AA807217.1 GI:2876793
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
           Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 517)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
           Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   On Jan 19, 1998 this sequence version replaced gi:2150641.
           Contact: Robert Strausberg, Ph.D.
           Tel: (301) 496-1550
           Email: Robert_Strausberg@nih.gov
           Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
           Ph.D. Gerald Marti, M.D.
           cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
           Bonaldo, Ph.D.
           cDNA Library Arrayed by: Greg Lennon, Ph.D.
           DNA Sequencing by: Washington University Genome Sequencing Center
           Clone distribution: NCI-CGAP clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           www-bio.llnl.gov/bbrp/image/image.html
           Insert Length: 1161   Std Error: 0.00
           Seq primer: -40m13 fwd. ET from Amersham
           High quality sequence stop: 500.
FEATURES   Location/Qualifiers
     source   1..517
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1351744"
                     /clone_lib="NCI_CGAP_GCB1"
                     /tissue_type="germinal center B cell"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from human tonsillar cells enriched for
                     germinal center B cells by flow sorting (CD20+, IgD-),
                     provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
                     (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
                     primed with a Not I - oligo(dT) primer
                     [5'-TGTTACCAATCTGAAGTGGGAGCGCGGCCGCTCATTTTTTTTTTTTTTTTTTT-
                     3']. Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT   177 a   102 c   113 g   125 t
ORIGIN

Query Match   20.3%;   Score 501.2; DB 38; Length 517;
Best Local Similarity   98.4%;   Pred. No. 2.7e-84;
Matches 506; Conservative   0; Mismatches   8; Indels   0; Gaps   0;

Qy 51 aggggacgcaccgggcgcgctcaacaaacctatggaagagtcagggtgtgtccccgccggtta 110
        | ||||| ||||||||| | |||| ||||||||| |||||| |||||||||| ||| ||||||
Db 4   ATGGACGGCCACCGGGCCTCGACAGACCTATCGAGAGATCAGGGTGTGCCCTCCCGGCGCCTTA 63

Qy 111 tcgggccaccaagctgtggaatgaagttaccacacatctttccgagaggaggaatgccctctaag 170
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db    64 TCGGGCCCAAGCTGTGGAATGAAGTACCACATCTTTCGAGCAGGAAATGCCTCTAAG 123
            ||||||||||||||||||||  ||||||||||||||||||||||||
Qy   171 aaaacacagacaacactttaaaaaatatgtgcaattgttcacgcagggagagcagtgga 230
            ||||||||||||||||||| |||||||||| ||||||||||||| |||||||||||||
Db   124 AGAACACAGACAACACTTTAAAAAATATGGCAATTGTTCACAGCAGGGAGAGCAGTGGA 183

Qy   231 ttggctttatgacctattaagaaataatagcaattttggtccctgaagttacaagcaaca 290
            ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   184 TTGGCTTTATGACCTATTAAGAAATAATAGCAATTTTGGTCCTGAAGTTACAAGCAACA 243

Qy   291 gactaccccaactgttgaggaaaatttcttaagaaaatcatgtaattgaagaatcaaaggag 350
            |||||| |||||||||||||||||||||||||||||| ||||| ||||||||||||||||
Db   244 GACTATCCACACTGTTGAGGAAAATTTCTTAAGAAATCATGTTAATTGAAGAATCAAAGGAG 303

Qy   351 gtgggggatcagaaaatgttgatgataacaacaacaagctcttcagaattccctgcaacttcgcc 410
            |||||||||||||| ||||||||||||||||||||||||||||||| |||||||||||||||
Db   304 GTGGGGATCAGAAAATGTTGATGATAACAACCAGCCTCTTCAGATTTCCCTGCAACTTCGCC 363

Qy   411 acttaaaactctaccacgaaggtatccagaattgagaagaaaacaacaacttgagaaacttttc 470
            ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   364 ACTTAAAACTCTACCACGAAGGTATCCAGAATTGAGAAAGACAACATAGAGAACTTTTC 423

Qy   471 caaagataaagatagacatttttaaattacgaaactctatctcgtagaaactcctaaaggca 530
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   424 CAAAGATAGAGATAGACATTTTTAGAATTACGAAACTTATCTCGTAGAAACTCCTAAAAGGCA 483

Qy   531 tggattacacatttatctcaggagaaaatggcgagaa 564
            ||||||||||||||||||||||||||||
Db   484 TGGATTACACATTTATCTCAGGAAAAATGGCGAGAA 517
```

RESULT 4
AI810084

```
LOCUS       AI810084           503 bp    mRNA            EST       07-JUL-1999
DEFINITION  wf65c09.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:2360464 3', mRNA sequence.
ACCESSION   AI810084
VERSION     AI810084.1  GI:5396650
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 503)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     On Feb 17, 1998 this sequence version replaced gi:2889728.
            Contact: Robert Strausberg, Ph.D.
            Email: Robert_Strausberg@nih.gov
            Tel: (301) 496-1550
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco
            High quality sequence stop: 457.
FEATURES             Location/Qualifiers
     source          1..503
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="Soares_NFL_T_GBC_S1"
                     /clone="IMAGE:2360464"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     This clone is available royalty-free through LLNL ; contact the
                     libraries (fetal lung NbHL19W, testis NHT, and B-cell
                     NCI_CGAP_GCB1) were mixed, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from pools of 5,000 clones made
                     from the same 3 libraries. The pools consisted of
                     I.M.A.G.E. clones 297480-302087, 682632-687229,
```

```
BASE COUNT    160 a    99 c    96 g   148 t
ORIGIN
```

Query Match 20.38%; Score 499.8; DB 61; Length 503;
Best Local Similarity 99.6%; Pred. No. 4.9e-84;
Matches 501; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy  1001 ctcattagaagaggggttgcccatcagagctctacagttatgttgttgttacttccccac 1060
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   503 CTCATTTAGAGAGGGGTTGCCCATCAGATCCTCTACACGTTATGTTGTTGTTACTTCCCCAC 444

Qy  1061 caaatcgtagaaaagcttcaacttttaatgctatgtgtattctccgaatgagtcaaaatgttg 1120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   443 CAAATCGTAGAAAAGCTTCAACTTTTAATGCCTATGATTTCCCGAATGAGTCAAAAATGTTG 384

Qy  1121 atatgcccaaacttcatgatggcaatgggtacgagagtcacctgaatgatgatacaacctttctc 1180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   383 ATATGCCCAAACTTCATGAATGGGTACGAGGGTCACTGAATGAGTCAACATACCTTTTTCTC 324

Qy  1181 gatgtgtgttatgctgtcgtgctgaagaagagtgtctagagcttcttgctcaggagagattag 1240
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   323 GATGTGTGTTATGCGTGTCGTGCTGAAGAAGTGGAGTCTTGATGAGCTTCTTGCTGGAAGATTAG 264

Qy  1241 tttcttcttaatggatcatcaggaaatctttcaagtacccctcttacttacagaactg 1300
            ||||||||||||||||||||||||||||||||||||||||||||
Db   263 TTTCTTTCTTAATGGATCATCACAGAAATCTTCAAGTACCCTCTTACTTACAGACTG 204

Qy  1361 tattttgctctttgccaactcatcatactactactaaaaagggacatatgcaattcctcggagagttgat 1420
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   143 TATTTGCTCCTTTGCCAACTTACTCATACTCATGGACATATGCAATTGAGG 84

Qy  1421 agcaaaaagttctctacctcaagctgcaattgcagaaacttttagaaaaatattataaa 1480
            ||||||||||||||||||||||||||||||||||||||||||||||
Db    83 AGCAAAAAAGTTTCTACCTCTCAAGCGCAATTGCAGAAAACTTTCAGGAACCA 24

Qy  1481 acaggagtttacctctaaaggag 1503
            |||||||||||||||||||
Db    23 ACAGGAGTTTACCTCTAAAGGAG 1
```

RESULT 5
AA545060

```
LOCUS       AA545060           692 bp    mRNA            EST       04-AUG-1997
DEFINITION  vj9ia10.s1 Knowles Solter mouse 2 cell Mus musculus cDNA clone
            IMAGE:944442 5', mRNA sequence.
ACCESSION   AA545060
VERSION     AA545060.1  GI:2306134
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 692)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The Washu-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     On May 5, 1995 this sequence version replaced gi:797839.
            Contact: Marra M/Mouse EST Project
            Washu-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
```

```
RESULT     6
           AA213817
LOCUS       AA213817/c      456 bp     mRNA             EST        13-AUG-1997
DEFINITION  zr91d11.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:683061 3',
            mRNA sequence.
ACCESSION   AA213817
VERSION     AA213817.1  GI:1812444
KEYWORDS    EST.
SOURCE      human.
```

```
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:541298
High quality sequence stop: 398.
FEATURES             Location/Qualifiers
source               1. .692
                     /organism="Mus musculus"
                     /strain="B6D2 F1/J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:944442"
                     /clone_lib="IMAGE:944442"
                     /clone_lib="Knowles Solter mouse 2 cell"
                     /tissue_type="embryo"
                     /dev_stage="2:cell"
                     /lab_host="DH10B"
                     /note="Organ: embryo; Vector: pBluescribe (modified);
                     Site_1: MluI; Site_2: SalI; Cloned unidirectionally from
                     mRNA prepared from 13,500 2-cell stage embryos, Primer:
                     SalI(dT): 5'-CCGTCGACGTGGACCGTTTTTTTTTTTTTT-3',
                     were cloned into the MluI/SalI sites of a modified
                     pBluescribe vector using commercial linkers (NEB).
                     Average insert size: 1.2 kb."
BASE COUNT      236 a    129 c    138 g    189 t
ORIGIN

Query Match      14.7%;    Score 362.6;   DB 35;   Length 692;
Best Local Similarity   86.4%;   Pred. No. 2.1e-58;
Matches 412;   Conservative 0;   Mismatches 64;   Indels 1;   Gaps 1;

Qy    79 atgggaggtcaggggtgtgtcctcccccggggccccaccaaagtcgttgaaagagtt 138
         |||||||||||| || ||||||||||| || |||||| ||||||||||||||||||
Db    53 ATGGAGGCGTCAGGGTGCTCTCTTGGACCTTACCGGGCCACCCAAGCTGTTGGAAGTA 112

Qy   139 accacatccttttcgagcagggaatgcctctaagaaaacacagacaacacttaaaaatat 198
         ||||||||||||| ||||||||||||||||||||||||||| ||| |||||||||||||
Db   113 ACCACATCTTTTCGAGTAGGAATGCCTTTAAGAAAACATAGACAACACATTTAAAAATAT 172

Qy   199 ggcaattgttcacagcaggagaagcagtggagttggcttatgacctattaagaaataat 258
         ||||||||||||||||||||||||||| ||||| || |||||||| ||||| ||||
Db   173 AGCAAATGTTTCACTCCAGTTGAAGCTATAGATTGGCTTTATGAATCGTTAAGAAGTAAT 232

Qy   259 agcaatttttggtcctcaagttacaagcaagacatatccaactgttgaggaaattctt 318
         ||||||||| |||||| ||| |||||||||| || |||||||| |||||||||||
Db   233 AGTAATTTTGGTCCCTGAAGTTACAAGCAACAGACTATTCAACTATTGAGAAAATTTCTT 292

Qy   319 aagaatcatgtaattgaagatatcaaaggggaggtgggggatcaagaaaatgttgatgataac 378
         |||| ||||||||||||||||||||||| || |||||||||| || |||||||
Db   293 AAGAATCATGTGATTGAAGATATCAAAGGTAGATGGGGATCGGGAAAGCCTTGATCGACAAC 352

Qy   379 aaccagctcctcagattctcctgccaacttcgcccactaaaaactctaccacggaagtatcca 438
         |||| ||||||||||| || ||||| |||||| || |||| |||| ||| ||||
Db   353 AATCAACTATTCAGAGATTTCCTGCAACTTCTCTCCAACTTTTCACAAAGACATACA 412

Qy   439 gaattgagaaaaaacaacatagagaacttttcccaaagataaagacatcatttttaaatta 498
         |||| |||||||| |||||||||| ||||||| ||||| ||||| |||||||||||||
Db   413 GAAATCAAGAAAAACCAACAACAGAGTACTTTCTCAAAGATGGTATATTTTTAAATTA 472

Qy   499 cgaaacttatctcgttagaacctcctaaaaaaggatcattggattacattttatctcaggaaaat 555
         ||||||||||||| |||||| ||||| |||| ||| |||||| |||||||| || ||||||||
Db   473 CGAAACTTATCTC-AAAAACTTCTAAGAAAGCGTGGATTACATTCTCTCAGGGAAAT 528
```

```
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
          Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1  (bases 1 to 456)
AUTHORS    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
           Tumor Gene Index
JOURNAL    Unpublished (1997)
COMMENT    On Apr 14, 1993 this sequence version replaced gi:315340.
           Contact: Robert Strausberg, Ph.D.
           Email: Robert_Strausberg@nih.gov
           Tel: (301) 496-1550
           This clone is available royalty-free through LLNL ; contact the
           IMAGE Consortium (info@image.llnl.gov) for further information.
           Insert Length: 1729    Std Error: 0.00
           Seq primer: -41m13 fwd. ET from Amersham
           High quality sequence stop: 317.
FEATURES             Location/Qualifiers
source               1. .456
                     /organism="Homo sapiens"
                     /db_xref="GDB:586381"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:683061"
                     /clone_lib="NCI_CGAP_GCB1"
                     /tissue_type="germinal center B cell"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from human tonsillar cells enriched for
                     germinal center B cells by flow sorting (CD20+, IgD-),
                     provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
                     (NCI) and Dr. Gerald Marti (CBER).  cDNA synthesis was
                     primed with a Not I - oligo(dT) primer
                     [5'-TGTTACCAATCTGAAGTGGAGGGCGCCGCCCTCATTTTTTTTTTTTTTTTT-
                     3']. Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT      141 a    88 c    139 g    139 t
ORIGIN

Query Match      14.5%;    Score 356.8;   DB 30;   Length 456;
Best Local Similarity   95.2%;   Pred. No. 2.7e-57;
Matches 432;   Conservative 0;   Mismatches 14;   Indels 8;   Gaps 6;

Qy  1054 ccccacacaaatcgtagaaagcttcaactttaatgcgtatgattt-cccgaatgagtca 1112
         || || |||||||||||||||||||| ||||||||| ||| |||||||||||||||
Db   452 CCCAACAAANTCGTAGAAGCTTCAAACTTTAATCGGTAGAAGTTCCCGAATGAGTCA 393

Qy  1113 aaat--gttgatatgcccaaactt-catgatgcaaatgggtacgaggtcaac-tgatgatac 1168
         ||||   |||||||||||||||| |||||||||||||||||||||||||| ||||||| |
Db   392 AAAATGGTGGATATGCCCAAACTTCCATGATGCAAATGGGTACGAGGTCACTTGATGATNC 333

Qy  1169 ataccttttctcgatgtgtgttatgtgtgtcgaagaagtggatcttgatgagcttctg 1228
         ||||||||| ||||| |||||||| || ||| |||||||||| |||||||||| || |
Db   332 ATACCTTTTTCTCGAATGTGTGTTTATGCCTGCGTAAGAAGTGGATNCTTGATGACTTCTTG 273

Qy  1229 ctgggaaagattagtttctttcttcaatggatcatcaggagaaattcttcaagtaccctctt 1288
         |||| ||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   272 CTGGAAGAATTAGTTTCTTTCTTCAATGGATCATCAGGAGAAATTCTTCAAGTACCCTCTT 213

Qy  1289 acttacagactgcagtggaaaaacatcttgactcactttaaaaaaggagacaatattgaaatc 1348
         ||||||||||||| |||||||||||||||||| |||||||||||||| ||| ||| |||| ||
Db   212 ACTTACAGACTGCGAGTGGAAAAACATCTTGACTACTTTAAAAAAGGGACATATATGAAAATC 153

Qy  1349 ctggagatggactatttgtctcctttgccaacttacctcagcagcagattagtgctc 1408
         ||||| |||||||||||| |||||||||||||||| ||| ||| |||||||||||||
Db   152 CTGGAGGATGGACTATTTGCTCCTTTGCCAACTTACTCAGTCATTAAGCAGATTAGTGCTC 93

Qy  1409 aggagtttgagtcaagcaaaaagtttctacctctcaa-gctgcaattgcagaaacttttgaa 1467
         ||||||||||| |||||||||||||||||||||||||  ||||||||||||||||||||||
```

```
Db      92  AGGAGTTTGATGAG--AAAAGTTTCTACCTCTCAATGCTGCAATTGCAGAACTTTTAGAA  35
            |||||||||||||||  |||||||||||||||||||||||||||||||||||||||||
Qy    1468  aatattattaaaaacaggagggtttaccttctaaagg                        1501

Db      34  AATATTATTAAAAACAGGAGTTTACCTCAAAAGG  1
```

RESULT 7

```
AA756999/c

LOCUS       AA756999              328 bp    mRNA             EST  29-DEC-1998
DEFINITION  ah54b09.s1 Soares_testis_NHT Homo sapiens cDNA clone 1309433 3',
            mRNA sequence.
ACCESSION   AA756999
VERSION     AA756999.1  GI:2804862
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

COMMENT

```
            NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
            Unpublished (1997)
            On Jan 19, 1998 this sequence version replaced gi:2151285.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
```

JOURNAL

FEATURES   Location/Qualifiers
     source   1..328
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="1309433"
                     /clone_lib="Soares_testis_NHT"
                     /sex="male"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from mRNA obtained from Clontech
                     Laboratories, Inc., and primed with a Not I - oligo(dT)
                     primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCGCCCCCCAATTTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization to Cot5, and was
                     constructed by Bento Soares and M. Fatima Bonaldo."

Insert Length: 570   Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 325.

```
Query Match             13.2%;    Score 326;  DB 38;    Length 328;
Best Local Similarity   100.0%;   Pred. No. 1.6e-51;
Matches 326;  Conservative    0;   Mismatches    0;   Indels   0;   Gaps   0;

BASE COUNT         83 a         55 c         49 g        141 t
ORIGIN
```

```
Qy    1477  aaaaacaggagttacctctaaaggagaaagaaagaaaaactaaaacagttcagaaggaa  1536
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     328  AAAAACAGGAGTTTACCTCTAAACGGAGAAAAGAAAAACTAAAACAGTTTCAGAAGGAA  269

Qy    1537  tatccttttgatatatcagaaaaagatttccaacacacggagttgaagcagccaactttttggt  1596
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     268  TATCCTTTTGATATATCAGAAAAAGATTTCCAACCACGGAGAGTTGAAGCAGCCACTTTTTGGT  209
```

RESULT 8

```
AA825936/c

LOCUS       AA825936              482 bp    mRNA             EST  09-JUN-1998
DEFINITION  od59h06.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1372283 3',
            mRNA sequence.
ACCESSION   AA825936
VERSION     AA825936.1  GI:2899248
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

COMMENT

```
            NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
            Unpublished (1997)
            On May 5, 1995 this sequence version replaced gi:797816.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
            Ph.D., Gerald Marti, M.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
```

JOURNAL

REFERENCE   1 (bases 1 to 482)
  AUTHORS   Bonaldo,M.F.
  TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index

FEATURES   Location/Qualifiers
     source   1..482
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1372283"
                     /clone_lib="NCI_CGAP_GCB1"
                     /tissue_type="germinal center B cell"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from human tonsillar cells enriched for

Insert Length: 701   Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 479.

```
Qy    1597  gacaaacctacaatcaagacaccaatgctgattttaagaagaaaaccaagttccgtagtcta  1656
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     208  GACAAACCTACAATCAAGCAACCAATGCTGATTTTAAGAAGAAACCAAGTTCCGTAGTCTA  149

Qy    1657  agataaactaactgaattaaaaatatgttgataatacttgtggaacttttgataatgaagccat  1716
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     148  AGATAACTAACTGAATTAAAAATATGTTGATAATACTTGTGGAACTTTTGATAATGAAGCCAT  89

Qy    1717  atctgagaatgtagctacctccaaaggagagtctgtcattaataaggtatttctaaataaac  1775
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      88  ATCTGAGAATGTAGCTACTCCAAAAGGAGAAGTCTGTCATTAATAAGGTATTTCTAAATAAAC  29

Qy    1777  acattatgtaaggaagtgccaaaata  1802
            ||||||||||||||||||||||||||
Db      28  ACATTATGTAAGGAAGTGCCAAAATA  3
```

RESULT 9

LOCUS   AQ376588   698 bp   DNA   GSS   29-JUN-1999
DEFINITION RPCI-11-161F2.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-161F2,
            genomic survey sequence.
ACCESSION   AQ376588
VERSION   AQ376588.1 GI:4347611
KEYWORDS   GSS.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 698)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE   Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT   Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Other_GSSs: RPCI-11-161F2.TV
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieterdejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from

BASE COUNT   144 a   73 c   97 g   168 t
ORIGIN

Query Match   13.2%; Score 325; DB 39; Length 482;
Best Local Similarity 100.0%; Pred. No. 2.3e-51;
Matches 325; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   664 atctacctgcaaacacattttaggtgtgtccatccctagagagaagtcataaaatccaaaacaa 723
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   482 ATCTACCTGCAAACACATTTTAGGTGTGTCCATCCCTAGAGAGAAGTCATAAAATCCAAAACAA 423

Qy   724 gtaattcccaatatatatattgtacaacatgggcccaatacaagtaaacgtggagttagttata 783
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   422 GTAATTCCCAATATATATATTGTACAACATGGGCCCAATACAAGTAAACGTGGAGTTAGTTATA 363

Qy   784 ctacaaaacaacaatcaagatcgaacctccctcactggtgattatctcgccatgaagtgcctagca 843
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   362 CTACAAAACAACAATCAGATGACCTCCCTCACTGGGTGATTATCTCGCCATGAAGTGCCTAGCA 303

Qy   844 aattggccaagaagcaatgatgatgaataatccaacttatgttgaacgagatgta 903
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   302 AATTGGCCAAGAAGCAATGATATGAATAATCCAACTTATCTTGGATTTGAACGAGAGATGTA 243

Qy   904 ttcaagaacaaatcgcagaatatattttctcagatctccctgaacctcactttttgaatat 963
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   242 TTCAGAACAATCGCAGAATTATTTCTCTAGATCTCCCTGAACCTCTACTTACTTTTGAATAT 183

Qy   964 tacgaattatttgtaaacatttgg 988
          ||||||||||||||||||||||||
Db   182 TACGAATTATTTGTAAACATTTGG 158

RESULT 10

LOCUS   AA420348   387 bp   mRNA   EST   16-OCT-1997
DEFINITION vd01g02.s1 Knowles Solter mouse 2 cell Mus musculus cDNA clone
            IMAGE:791282 5', mRNA sequence.
ACCESSION   AA420348
VERSION   AA420348.1 GI:2080825
KEYWORDS   EST.
SOURCE   house mouse.
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 387)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE   The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT   On Sep 12, 1996 this sequence version replaced gi:1394414.
            Contact: Marra M/Mouse EST Project
            Washington University School of MedicineP
            WashU-HHMI Mouse EST Project
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

BASE COUNT   214 a   129 c   146 g   209 t
ORIGIN

Query Match   12.7%; Score 312.4; DB 102; Length 698;
Best Local Similarity 96.7%; Pred. No. 5.1e-49;
Matches 319; Conservative   0; Mismatches 11; Indels   0; Gaps   0;

Qy   989 gcttgctgcaacctcattttagagaggagggttgccatcgcatctctaccagttatgttgttgt 1048
          |||||| |||||||||||||||||||||||||||||||||||||||| |||||| ||||||||||
Db   369 GTTTGCTGCAACCTCATTTAGCAGGAGGGTTGCCATCGACATGTCTACAGTTATGTTGTTTTGT 428

Qy   1049 tacttcccccaccaaatcgtagaaaagcttcaacttttaaatgcgtatgtatgaattcccgaatga 1108
          || |||||| ||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db   429 TATTTCCCCCACCAAATCATAGAAAAGCTTCAACTTTTAAATGGCGTATGTATGAATTCCCGAATGA 488

Qy   1109 gtcaaaagtgttgattatgcccaaaactcatgatgcaaatgggttactgagggtcactgatgatac 1168
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   489 GTCAAAAGTGTTGATTATGCCCAAAACTTCATGATGCAAATGCAATGGGGTCGAGATCACTGTTGATAC 548

Qy   1169 ataccttttctcgatgtgttatgtgctgtgctgaagaagtggaatcttgatgagcttcttg 1228
          ||||| |||||||||||||||||||||||||||||| ||||||| ||||||||||||||||
Db   549 ATACCGTTTCTCAAATGTGTTATCTGCTGCTGCTGCGGAAGAAGTGGACCTTGATGAGCTTCTTG 608

Qy   1229 ctggaagattagtttcttcttttcttaaatggatcatcaggaaatttcttcaagtaccctctt 1288
          ||||||||||||||||||||||| |||||||||||||||||||||||||||| ||||||||||
Db   609 CTGGAAGATTAGTTTCTTCTTTTCTTAATGGATCATGATCAGGAAAATTCTTCAAGTACCCCTCTT 668

Qy   1289 acttacaagactcgcagtggaaaacatcttg 1318
          ||||||||||||||||||||||||||||| ||
Db   669 ACTTGCAGACTCGCAGTGGAAAAACATCTTG 698

Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:475690
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 375.

FEATURES   Location/Qualifiers
source   1..387
                /organism="Mus musculus"
                /strain="B6D2 F1/J"
                /db_xref="taxon:10090"
                /clone="IMAGE:791282"
                /clone_lib="Knowles Solter mouse 2 cell"
                /note="Organ: embryo; Vector: pBluescribe (modified);
                Site_1: MluI; Site_2: SalI; Cloned unidirectionally from
                mRNA prepared from 13,500 2-cell stage embryos; Primer:
                SalI(dT): 5'-CGGTCGACGTCGACCGTTTTTTTTTTTTTTTT-3'; cDNAs
                were cloned into the MluI/SalI sites of a modified
                pBluescribe vector using commercial linkers (NEB)."
                /lab_host="DH10B"
                /dev_stage="2-cell"
                /tissue_type="embryo"

BASE COUNT   122 a   79 c   79 g   107 t
ORIGIN

Query Match   11.0%;   Score 271; DB 33;   Length 387;
Best Local Similarity   88.1%;   Pred. No. 3e-41;
Matches   295;   Conservative   0;   Mismatches   40;   Indels   0;   Gaps   0;

Qy   79 atggagagtcaggttgtgcctcccggggccttatcggggcaccaagctgtggaatgaagtt 138
          ||||||||| || | ||| ||||||||||||||||||||||||||| ||||||||||
Db   53 ATGGAGAGGTCAGGGTGCTCCCTTGGACCTTACGGCGCCACCAAGCTGTGGAATGAAGTA 112

Qy   139 accacatcatcttttcgagaggaatgccttcctaagaaaaacaacagacaaacacttaaaaatat 198
          | ||||| |||||||||||||||||| |||||||| | || ||||| ||| |||||||||||
Db   113 ACCACATCTTTCGAGTAGGAATGCCTTAAGGAAAAACATAGACAACAACAACATTTAAAAAATAT 172

Qy   199 ggcaattgttcacagcaggggagaaagcagtgatgattggctcttattgtgaccattaagaaataat 258
          |||||||||| |||||||||||| | ||||| ||| ||||||||||||||||| ||||| || |
Db   173 AGCAATTGTTTCACTCGCAGTTGGAAGCTATAGAGATTGGCTTTATGATCTGTTGTTAAGAAGTAAT 232

Qy   259 agcaatttttggtccctgaagttacaaggcaacaagactatccaactgttggaggaaattcctt 318
          || |||| |||||||||| || ||||||| | |||| ||| || |||||| ||| ||||||
Db   233 AGTAATTTTGGTCCCTGAGGTTACAAGACACAACAGACTATTCAACTATTGAGAAAATTTCTT 292

Qy   319 aagaatacatgtaattgaagagaatatcaaggggaggagaatgtgtggaaaatgttgatgataac 378
          |||||||||| |||||||||||| ||||| ||||||||||||| || | || ||||| ||||
Db   293 AAGAATCATGTGATTGAAGAGATATCAAAGGGATAGGGGGATCGGAAAGCCTTGATGACAAC 352

Qy   379 aaccagtcttccagatttcctgcaacttcgcacact 413
          |||| |||||||||||||||||||||||||| ||||
Db   353 AATCAACTATTCAGATTTCCTGCAACTTCCCACT 387

RESULT 11
LOCUS   AA907054   361 bp   mRNA   EST   10-JUN-1998
DEFINITION omlg08.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:1540670 3', mRNA sequence.
ACCESSION   AA907054
VERSION   AA907054.1 GI:3042514
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 361)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

---

JOURNAL   Tumor Gene Index
COMMENT   Unpublished (1997)
            On Jan 14, 1998 this sequence version replaced gi:1797337.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1289   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 350.

FEATURES   Location/Qualifiers
source   1..361
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:1540670"
                /clone_lib="Soares_NFL_T_GBC_S1"
                /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                Equal amounts of plasmid DNA from three normalized
                libraries (fetal lung NbHL19W, testis NHT, and B-cell
                NCI_CGAP_GCB1) were mixed, and ss circles were made in
                vitro. Following HAP purification, this DNA was used as
                tracer in a subtractive hybridization reaction. The driver
                was PCR-amplified cDNAs from pools of 5,000 clones made
                from the same 3 libraries. The pools consisted of
                I.M.A.G.E. clones 297480-302087, 682632-687239,
                726408-728711, and 729096-731399. Subtraction by Bento
                Soares and M. Fatima Bonaldo."
                /lab_host="DH10B"

BASE COUNT   128 a   71 c   78 g   84 t
ORIGIN

Query Match   10.9%;   Score 269.4; DB 39;   Length 361;
Best Local Similarity   99.6%;   Pred. No. 6e-41;
Matches   270;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

Qy   390 cagaatttcctgcaacttcgccacttcaaaactctaccacgaaggtatccagaagaattgagaa 449
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   91 CTGAATTCCTGCAACTTCGCCCACTTAAAACTCTACCACGAAGGTATCCAGAATTGAGAAA 150

Qy   450 aaacaacatagaagaacttcttcccaaaagatataagatatagacattcttttaaattacgaaacttatc 509
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   151 AAACAACATAGAGAACTTTTCCAAAGATAAAGATAGCATTTTTAAAATAACGAAACTTATC 210

Qy   510 tcgttagaactcctaaaaggcatggtaatacacattatctccagggaaatctggcgtgaaaaataaa 569
          ||||||||||||||||||||||||||||||| |||||||||||||||||||| |||||||||
Db   211 TCGTTAGAACTCCTAAAAGGCATGGATTACATTTATCTCCAGGGAAATGGCGAGAAAAATAAA 270

Qy   570 gcatgaaatatcaaatgaagatcaagaagaaaatgcaaatctgataataagaagaactaagccagga 629
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   271 GCATGAAATAATCAATCAATGAAGATCAAGAAGAAAATCGAAATCTGATAATTAGAGAAGACTAAGCCAGGA 330

Qy   630 agatgttgaagaagttggagatatgttatt 660
          |||||||||||||||||||||||||||
Db   331 AGATGTTGAAGAAGTTGGAGATATAGTTATT 361

RESULT 12
AQ261251/c
LOCUS   AQ261251   378 bp   DNA   GSS   24-OCT-1998
DEFINITION CITBI-E1-2506G1.TR CITBI-E1 Homo sapiens genomic clone 2506G1,
            genomic survey sequence.
ACCESSION   AQ261251
VERSION   AQ261251.1 GI:3787775
KEYWORDS   GSS.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 378)
AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,

RESULT 13

```
LOCUS       AQ791096       496 bp       DNA       GSS       03-AUG-1999
DEFINITION  HS_2189_B1_F05_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2189 Col=9 Row=L, genomic survey
            sequence.
ACCESSION   AQ791096
VERSION     AQ791096.1  GI:5698720
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 496)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
```

Query Match   10.5%; Score 258.2; DB 100; Length 378;
Best Local Similarity 93.7%; Pred. No. 7.4e-39;
Matches 269; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```
Qy  702 agaagtccataaatccaaaacaagtaattccccaatatattatatacaacaatgccaatac 761
        ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db  377 AGAAGTCATAAAATCCAAAATCAGTAATTCCCTATATATATAATGTACAACAATGCCAATAC 318

Qy  762 aagtaaacgtggagtagttatactacaaaaacaaaataagatgacctccctcactggtatt 821
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  317 AAGTAAACATGGAGTAGTTACACTACAAAATCAAATCAGATGACCTCCCTCACTGGGTATT 258

Qy  822 atctgccatgaagtgcctagcaaattggcccaagagaagcaatgatatgaataatccaactta 881
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  257 ATCTGCCATGAAGTGCCTAGCAAATTGGCCAAGAAGTGATGATAACGAATAATCCGACTTA 198

Qy  882 tgttggaatttgaacgaggatgtattcagcagaacaatgcagaattatttttctagatctccctga 941
        ||||||||||||| |||||||||||||||||||||||||||||||||| |||
Db  197 TGTTGGAATTTGAACAAGAATATATTCAGCACAACATTACACAGATTAATTTCTAGATCTCCTTGA 138

Qy  942 acctctacttacttttgaatattacgaaatatttgtaaacactttgg 988
        |||||||||||||||||||||||||||||||||| |||
Db  137 ACCTCTCCTTACTTTGAATATTATGAATTATTTTGTGAACATTTTGG 91
```

COMMENT   The Institute for Genomic Research
            Department of Eukaryotic Genomics
            Contact: Mark Adams
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.

BASE COUNT   119 a   61 c   71 g   127 t
ORIGIN
```
FEATURES             Location/Qualifiers
     source          1..378
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="2506G1"
                     /clone_lib="CITBI-E1"
                     /sex="male"
                     /cell_type="sperm"
                     /note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI;
```

TITLE   Use of a random human BAC End Sequence Database for Sequence-Ready
            Map Building
JOURNAL   Unpublished (1998)
COMMENT   Other_GSSs: CITBI-E1-2506G1.TF

AUTHORS   Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
            Venter,J.C.

```
LOCUS       AQ791096
DEFINITION  HS_2189_B1_F05_MR CIT Approved Human Genomic Sperm Library D
ACCESSION   AQ791096
VERSION     AQ791096.1  GI:5698720
KEYWORDS    GSS.
SOURCE      human.
```

RESULT 14

```
LOCUS       AA240712       501 bp       mRNA       EST       12-MAR-1997
DEFINITION  mv30b1l.r1 GuayWoodford Beier mouse kidney day 0 Mus musculus cDNA
            clone IMAGE:656541 5', mRNA sequence.
ACCESSION   AA240712
VERSION     AA240712.1  GI:1865087
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 501)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
```

Query Match   8.4%; Score 206.2; DB 87; Length 496;
Best Local Similarity 91.2%; Pred. No. 3.8e-29;
Matches 229; Conservative 0; Mismatches 20; Indels 2; Gaps 1;

```
Qy 1122 tatgcccaaacttcatgatgcaatgggtacggaggtcactgagtgaatacacatacctttctcg 1181
        ||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   83 TATGTCCAAACTTCATGATGCAATGGGCCATGAATCACCGATGAATACATACCTTTTCTCG 142

Qy 1182 atgtgttgttatgctgtgctggagagaagtggaatctttgatctttgatgagcttcttgctggaagatttagt 1241
        ||| ||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db  143 ATGTGTTGCTATGCTGTGCTGGAGAAAAGTGGATCTTGATAAGCTTCTTGCTGGAGATTAGT 202

Qy 1242 ttctttcttaattggatcatcatcaggaaaaattcttcaagtacccctccttacttacagactgc 1301
        ||||| |||||||||| |||||||| |||||||||||||||||||||||| |||||
Db  203 NTCTTTCTTACTGGATCANTCATCAGGAAAAATTCTTCAAGTACCAACTTRATTTACAGACTGC 262

Qy 1302 agtggaaaaacatcttgactacttaaaaaaaggaacatattgaaaatcctggagaatggact 1361
        ||||| |||||||||| ||||||||||||||||||||||||||||||||| |||||||||
Db  263 AGTGGGAAAAACATCTTGAATACTTAAAAAAAGGGACAATATTGAAAAATC--TGAGATGGACT 320

Qy 1362 atttgctcctt 1372
        |||| |||||
Db  321 ATNGTTTCTTT 331
```

COMMENT   On Sep 10, 1998 this sequence version replaced gi:3552083.
            Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones may be purchased from Research Genetics (info@resgen.com).
            BAC end Web Server: http://www.htsc.washington.edu
            Plate: 2189 row: L column: 9
            Seq primer: M13 Reverse
            Class: BAC ends.
            High quality sequence stop: 496.

BASE COUNT   159 a   91 c   91 g   150 t   5 others
ORIGIN
```
FEATURES             Location/Qualifiers
     source          1..496
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=2189 Col=9 Row=L"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /sex="male"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
```

TITLE   Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
JOURNAL   Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE   99380589
AUTHORS   Hood,L.

Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R., and
Waterston,R.
TITLE   The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1402298.
        Contact: Marra M/Mouse EST Project
        WashU-HHMI Mouse EST Project
        Washington University School of Medicine
        4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
        Tel: 314 286 1800
        Fax: 314 286 1810
        Email: mouseest@watson.wustl.edu
        This clone is available royalty-free through LLNL ; contact the
        IMAGE Consortium (info@image.llnl.gov) for further information.
        MGI:1402389
        Seq primer: -28m13 rev1 ET from Amersham
        High quality sequence stop: 437.
FEATURES
source
        1..501
        /organism="Mus musculus"
        /strain="C57BL/6J"
        /db_xref="taxon:10090"
        /clone="IMAGE:655541"
        /clone_lib="GuayWoodford Beier mouse kidney day 0"
        /tissue_type="kidney"
        /dev_stage="newborn (day 0)"
        /lab_host="SOUR (kanamycin resistant)"
        /note="Organ: kidney; Vector:pBluescript SK-; Site_1:
        EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
        Oligo dT. Average insert size: 1.0 kb; Uni-ZAP XR Vector;
        -5' adaptor sequence: 5' GAATTCGGCACGAG 3' -3' adaptor
        sequence: 5' CTCGAGTTTTTTTTTTTTTTTTTT 3' Library provided
        Lisa Guay-Woodford."
BASE COUNT   125 a   115 c   136 g   125 t
ORIGIN

Query Match
Best Local Similarity   62.0%;   Score 185; DB 30;   Length 501;
Matches 313; Conservative 0; Mismatches 185; Indels 7; Gaps 1;   Pred. No. 3.5e-25;

Qy 779 ttactacaaaacaaatcagatgacctccctcactggtgtattatctgccatgaagtgcc 838
        |||||||| ||||| ||||| | |||||||||||||||||||||||||||||||||||||
Db   1 TCATCCTGGACGACAAGTGAAAGAACTTCCTCACTGGGTCTGCTGTCGTATGAAGTGTT   60

Qy 839 tagcaaaattggccaagaagcaaatgatatgaataatccaactatatgttggattgaacgag 898
        |||||||||||| |||| | |||||||| || |||| ||||||||||||||||||||||||||
Db   61 TGGCAAAATTGGCCCAACTGTACAGACTGGAAGCAGCCTATGTGACTGGGATTTGAGAAGG 120

Qy 899 atgtattcagaaacaatcgcagaattatttttctagatctccctgaacctctacttactttg 958
        ||||| ||| ||| ||| | |||||| ||||||||||| || |||||||| ||||||| |||
Db 121 ATGTCTTTAAGACTATAGACAGATTAACTATAGGTCACCTGAGGAGGCCACTCGCTACATTTC 180

Qy 959 aaatattaccgaattattgtaaacatttgggctgtcgctgcaactccatttagagaggggttg 1018
        ||| ||| |||||||||| ||| ||||| |||||||| |||||||||| ||||||||||||||
Db 181 ATCTCCTTTGATGCTTTCGTCCAGCGTCGTTTAGGTTGTCCTGC------CGAAGAGAAAACGG 233

Qy 1019 ccatcgatgctcacagttcatgtgttgttgttacttcccccaccaaaatcgtagaaagcttc 1078
        ||| ||| |||| ||| ||||| ||||||||||||||||||||| |||||||||||||||||
Db 234 CCAATGAACGACATTTCAGCTTTGCTTCGTCCTCCTCCCCTGAAAACAGAGAAAGTTAC   293

Qy 1079 aacttttaatgcgtatgattcccgaatgagagtcaaaaatggttgatgatgcccaaaacttcatg 1138
        || ||||| ||| |||||||| |||| | ||| |||| |||||||| || ||||||| |||||||
Db 294 AGCTGTTGATGAGGATGAGGCGAGGATCTGCCCTAAACAAGAGAGATGCCCCACTGAGTG   353

Qy 1139 atgcaatgggtaccagagtcactgatgatacataccttttctctcgatgtgtgttatgctgtg 1198
        ||| | || |||| || |||||| |||||| ||||||||||| || |||||| || ||| |||
Db 354 ATGGCTTCGGCACCCGCACACTGATAGGGTTCAGAACCTTTTCCCGGTTCGATCCCTGCCTCTA 413

Qy 1199 ctgaagaagtggatcttgatgagctcttctgctggaagattagttactacttagatgatgatc 1258
        ||||| ||| || |||| |||||| || | ||||||||||| ||||| ||| ||||||| |||
Db 414 AGGATGAAGTAGAGACTTAGATGAGTTACTAGCTGTCGCTGATTGGTGACATTCCTGATGATGGACA 473

RESULT 15
AQ415970

Qy 1259 atcatcaaggaaattcctcaagtacc 1283
        ||| ||||| || ||||| ||||| ||
Db 474 ATTACCAGGAGGATTCTGAAAGTGCC 498

LOCUS   AQ415970   564 bp   DNA   GSS   23-MAR-1999
DEFINITION RPCI-11-207H10.TJ RPCI-11 Homo sapiens genomic clone
            RPCI-11-207H10, genomic survey sequence.
ACCESSION   AQ415970
VERSION   AQ415970.1 GI:4474939
KEYWORDS   GSS.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 564)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE   Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT   Other_GSSs: RPCI-11-207H10.TV
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@djong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genet cs (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
            Seq primer: SP6
            Class: BAC ends.
FEATURES
source
            1..564
            /organism="Homo sapiens"
            /db_xref="GDB:7579281"
            /db_xref="taxon:9606"
            /clone="RPCI-11-207H10"
            /clone_lib="RPCI-11"
            /sex="Male"
            /cell_type="Lymphocytes"
            /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
            RPCI11 Human Male BAC Library"
BASE COUNT   179 a   106 c   116 g   161 t   2 others
ORIGIN

Query Match
Best Local Similarity   92.4%;   Score 174.2; DB 102;   Length 564;
Matches 182; Conservative 0; Mismatches 15; Indels 0; Gaps 0;   Pred. No. 3.6e-23;

Qy 989 gcttggctgcaacccattaggagagggggtggccacagatgctcacagtatgtgttgt 1048
        |||||| | |||||| || ||||||||| || |||| |||||||||| |||| ||||||
Db 368 GTTTGCTGCTGCAACCTCATTTAGGAGAGAGGGGTTCCCATCGAATTGTGTCGTTTGT 427

Qy 1049 tacttccccaccaaaatcgtagaaagctcaactctttaatgcgtatgattcccgaatga 1108
        |||||||| |||||||| | |||||||| ||||| ||||||||||||||| ||| |||||
Db 428 TACTTCCGCGCCACCAAAATTGTTAGAAAGCTTCAACTTTTAAATGCGTATGATTTCCAAATGA 487

Qy 1109 gtcaaaatgttgatatgcccaaaacttcatgatgatcaatggtacgaggtcactgatgatac 1168
        |||||| ||||||||| | ||||||||||||||||| ||||| || ||| |||||||||||||
Db 488 GTCAAAATGTTGATATGTCCAAAACTTCATGATGATGCATGANATACCGATGATAC   547

Qy 1169 atacttttctctcgatgt 1185
        |||| | |||||||||||
Db 548 ATACCTNTTCTCGATGT 564

Search completed: March 23, 2000, 21:52:11
Job time: 8155 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 24, 2000, 13:19:52 ; Search time 172.6 Seconds
                       (without alignments)
                       1385.209 Million cell updates/sec

Title: 1997
Perfect score: US-09-183-789-38_COPY_1_1997
Sequence: 1 ctcactataggagaccac............tctctgttagtaaagctgc 1997

Scoring table:
                       IDENTITY_NUC
                       Gapop 10.0 , Gapext 1.0

ched: 214294 seqs, 59861574 residues

Total number of hits satisfying chosen parameters: 428588

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database :   Issued_Patents_NA:*
               1:/cgn2_6/ptodata/1/ina/5A_COMB.seq:*
               2:/cgn2_6/ptodata/1/ina/5B_COMB.seq:*
               3:/cgn2_6/ptodata/1/ina/5C_COMB.seq:*
               4:/cgn2_6/ptodata/1/ina/5D_COMB.seq:*
               5:/cgn2_6/ptodata/1/ina/6_COMB.seq:*
               6:/cgn2_6/ptodata/1/ina/PCTUS9_COMB.seq:*
               7:/cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB ID	Description
1	69.6	3.5	6420	3	US-08-374-483-1	Sequence 1, Appli
2	66.8	3.3	4525	3	US-08-613-861-2	Sequence 2, Appli
3	59.6	3.0	1123	1	US-08-458-023B-3	Sequence 3, Appli
4	59.4	3.0	7218	1	US-08-232-463-14	Sequence 14, Appl
5	56.2	2.8	1700	4	US-08-839-581A-1	Sequence 1, Appl
6	52.6	2.6	9592	1	US-08-393-734-3	Sequence 3, Appl
7	52.2	2.6	120	3	US-08-454-557C-76	Sequence 76, Appl
8	52.2	2.6	120	3	US-08-340-426D-76	Sequence 76, Appl
9	52.2	2.6	120	4	US-08-450-673C-76	Sequence 76, Appl
10	52.2	2.6	120	6	PCT-US95-17111A-76	Sequence 76, Appl
11	52.2	2.6	3984	2	US-08-040-753-1	Sequence 1, Appl
12	51.4	2.6	9837	2	US-08-832-883-68	Sequence 68, Appl
13	51.4	2.6	9837	2	US-08-832-877-68	Sequence 68, Appl
14	50	2.5	4059	3	US-08-252-493C-4	Sequence 4, Appl
15	50	2.5	4059	6	PCT-US95-07554-4	Sequence 4, Appl
16	50	2.5	4059	6	PCT-US96-05611A-11	Sequence 11, Appl
17	48.8	2.4	1982	4	US-08-900-148-1	Sequence 1, Appl
18	48.2	2.4	404	2	US-08-594-031-154	Sequence 154, App
19	47.2	2.3	1672	1	US-08-172-331B-13	Sequence 13, Appl
20	46.8	2.3	115	3	US-08-454-557C-75	Sequence 75, Appl
21	46.8	2.3	115	4	US-08-340-426D-75	Sequence 75, Appl
22	46.8	2.3	115	4	US-08-450-673C-75	Sequence 75, Appl
23	46.8	2.3	115	6	PCT-US95-17111A-75	Sequence 75, Appl
24	42.6	2.1	8299	2	US-08-462-014-2	Sequence 2, Appli
25	40.4	2.0	1425	1	US-07-715-184-3	Sequence 3, Appli
26	40.4	2.0	1425	1	US-07-876-280-6	Sequence 6, Appli
27	40.4	2.0	1425	1	US-07-876-280-27	Sequence 27, Appl
28	40.4	2.0	1425	1	US-07-935-310A-1	Sequence 1, Appli
29	40.4	2.0	1425	1	US-07-828-788A-1	Sequence 1, Appli
30	40.4	2.0	1425	1	US-08-049-783-3	Sequence 3, Appli
31	40.4	2.0	1425	1	US-08-147-189-1	Sequence 1, Appli
32	40.4	2.0	1425	1	US-08-316-301A-7	Sequence 7, Appli
33	40.4	2.0	1425	3	US-08-904-278-1	Sequence 1, Appli
34	40.4	2.0	1425	6	PCT-US92-03624-7	Sequence 7, Appli
35	40.4	2.0	1425	6	PCT-US92-11337-1	Sequence 1, Appli
36	40.4	2.0	1425	6	PCT-US93-07409-1	Sequence 1, Appli
37	36.4	1.8	3093	1	US-08-252-956B-19	Sequence 19, Appl
38	36	1.8	741	2	US-08-616-368A-2	Sequence 2, Appli
39	34.8	1.7	1405	2	US-07-710-361-3	Sequence 3, Appli
40	34.8	1.7	2210	1	US-07-710-361-2	Sequence 2, Appli
41	34.6	1.7	5183	2	US-08-459-568-3	Sequence 3, Appli
42	34.6	1.7	5183	2	US-08-399-411-3	Sequence 3, Appli
43	34	1.7	201	4	US-08-930-274-25	Sequence 25, Appl
44	33.8	1.7	5181	2	US-08-257-073-10	Sequence 10, Appl
45	32.8	1.6	2121	4	US-08-897-340-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-08-374-483-1
; Sequence 1, Application US/08374483
; Patent No. 5880102
; GENERAL INFORMATION:
; APPLICANT: GEORGE, SAMUEL E.
; APPLICANT: BLAZING, MICHAEL A.
; TITLE OF INVENTION: ADENOVIRAL VECTOR SYSTEM
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/374,483
; FILING DATE: 17-JAN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 1579-83
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; TELEX: 207797 NIXN UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6420 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-374-483-1

Query Match   3.5%; Score 69.6; DB 3; Length 6420;
Best Local Similarity 77.8%; Pred. No. 1.1e-10;
Matches   84; Conservative   0; Mismatches   24; Indels   0; Gaps   0;

Qy   1 ctcactataggagaccac-cttggttaccgagctcggatccactagttaacggcccggccag   60
             ||||||||| ||| ||| | |||||| ||||||| |||||| |||||||| ||||| ||||
Db   975 CTCACTATAGGGAGAGACCCAAGCTTGGTTGTTCTTGGGCCGAGCTCGGATCCACTAGTTAACGGCCGCGCCAG 1034

```
Qy    61 tgtgctggaaagtgttcaaccagtgattattttattgacagcaactgg 108
         ||||||||||| ||| ||||| || ||||| ||||| || ||| |||
Db  1035 TGTGCTGGAATTAATTCTCGAGATATCCATCACACTGGCGCGGCCCTCG 1082
```

RESULT 2
US-08-613-861-2
; Sequence 2, Application US/08613861
; Patent No. 5843770
; GENERAL INFORMATION:
;   APPLICANT: Ill, Charles R. et al.
;   TITLE OF INVENTION: Antisense Constructs Directed Against Viral Post-Transcriptio
;   NUMBER OF SEQUENCES: 2
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: LAHIVE & COCKFIELD
;   STREET: 60 State Street, Suite 510
;   CITY: Boston
;   STATE: Massachusetts
;   COUNTRY: USA
;   ZIP: 02109-1875
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/613,861
;   FILING DATE: 13-APR-1994
;   CLASSIFICATION: 514
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/111,111
;   FILING DATE: 12-DEC-1909
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Attorney, Name Init
;   REGISTRATION NUMBER: 000000
;   REFERENCE/DOCKET NUMBER: oe
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (617)227-7400
;   TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 4525 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
US-08-613-861-2

Query Match   3.3%; Score 66.8; DB 3; Length 4525;
Best Local Similarity   97.1%; Pred. No. 6e-10;
Matches   68; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

```
Qy    1 ctcactataggagaccacgcttggttaccgagctcggatccacactagtaacggccgccag 60
        |||||||||||||||||||||||||||||||||||| ||||||||| ||||||||||||||
Db 3070 CTCACTATAGGAGAGACCCAAGCTTCGTTACCGAGCTCGGATCCACTAGTAACGGCCGACAG 3129
```

```
Qy   61 tgtgctggaa 70
        ||| |||||||
Db 3130 TGTGCTGGAA 3139
```

RESULT 3
US-08-458-023B-3
; Sequence 3, Application US/08458023B
; Patent No. 5667990
; GENERAL INFORMATION:
;   APPLICANT: Berka, Randy M.
;   APPLICANT: Yoder, Wendy
;   APPLICANT: Takagi, Shinobu
;   APPLICANT: Boominathan, Karuppan C.
;   TITLE OF INVENTION: ASPERGILLUS EXPRESSION SYSTEM
```

---

; NUMBER OF SEQUENCES: 16
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: NO. 5667990 No. 5667990disk of No. 5667990th America, Inc.
;     STREET: 405 Lexington Avenue
;     CITY: New York
;     STATE: New York
;     COUNTRY: USA
;     ZIP: 10174-6201
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: Patentin Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/458,023B
;     FILING DATE: 01-JUN-1995
;     CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Lowney Dr., Karen A.
;     REGISTRATION NUMBER: 31,274
;     REFERENCE/DOCKET NUMBER: 4086.010-US
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 212-867-0123
;     TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1123 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;     MOLECULE TYPE: DNA (genomic)
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
;   FRAGMENT TYPE: internal
;   ORIGINAL SOURCE:
;     ORGANISM: Humicola insolens
;     INDIVIDUAL ISOLATE: DSM 6995
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 126..806
US-08-458-023B-3

Query Match             3.0%;  Score 59.6; DB 1;  Length 1123;
Best Local Similarity   87.8%;  Pred. No. 3.7e-08;
Matches   65; Conservative   0; Mismatches   9; Indels   0; Gaps   0;

```
Qy    1 ctcactataggagaccacgcttggttaccgagctcggatccacactagtaacggccgccag 60
        |||||||||||||||||||||| ||||| |||||||||||||||||||||||| |||||||
Db    8 CTCACTATAGGAGAATATTAAGCTTCGGTACCGAGCTCGGATCCACTAGTAACGGCCGCCAG 67
```

```
Qy   61 tgtgctggaaagtg 74
        |||||||||| |||
Db   68 TGTGCTCTAAAGCG 81
```

RESULT 4
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
;   APPLICANT: DORNER, F.
;   APPLICANT: SCHEIFLINGER, F.
;   APPLICANT: FALKNER, F. G.
;   TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
;   NUMBER OF SEQUENCES: 52
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Foley & Lardner
;   STREET: 1800 Diagonal Road, Suite 500
;   CITY: Alexandria
;   STATE: VA
;   COUNTRY: USA
;   ZIP: 22313-0299

COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/232,463
: FILING DATE:
PRIOR APPLICATION DATA:
: CLASSIFICATION: 435
: APPLICATION NUMBER: US/07/935,313
: FILING DATE: 26-AUG-1991
: APPLICATION NUMBER: EP 91 114 300.6
: FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
: NAME: BENT, Stephen A.
: REGISTRATION NUMBER: 29,768
: REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703)836-9300
: TELEFAX: (703)683-4109
: TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
: LENGTH: 7218 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: CLONE: pTzgpt-F1s
US-08-232-463-14

Query Match
Best Local Similarity 5.3%; Score 59.4; DB 1; Length 7218;
Matches 21; Conservative 220; Mismatches 156; Indels 0; Gaps 0;

3.0%; Pred. No. 1.2e-07;

Qy 1495 tgtctctgcaagtgaccacacaagttacaagttatgtcttcggtgggaaactgtgccaaatac 1554
|||||| |||||| ||| || ||| || ||
Db 1474 TATCTATGCAAGTAGTTAAAGAGATAGAGAAATTGGTGTACRRRRRRRRRRRRRRRRRRRR 1415
Qy 1555 accacacaaatatctcctgccatggcaaaaaaattaatgatgtatataaaatatcaattaat 1614
RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR
Db 1414 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1355
Qy 1615 gaaagaagttcgaagtgtttgggcaaaaatatgaaagaaattttcaattttgcttgaaggt 1674
RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR
Db 1354 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1295
Qy 1675 acaaggatctatgaaagtcaaagaacaattgttgattaccatcaaggagagcagcaag 1734
:::
Db 1294 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1235
Qy 1735 gtttaaaaaagttgtcttaattcagcaactcgagaaggcgcttaaagaaataagattccca 1794
:::
Db 1234 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1175
Qy 1795 ctgccatctcagaaaagttaagccacatagagaaaaagatataattgtgttagtgtcaaagacca 1854
: ::: :
Db 1174 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1115
Qy 1855 aggagaaacaaggacatatgtctgtaggatggaacagg 1891
:::::::::::::::::::::::::::::::::::::
Db 1114 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1078

RESULT 5
US-08-839-581A-1/c
: Sequence 1, Application US/08839581A
: Patent No. 5958705
: GENERAL INFORMATION:
: APPLICANT: Staunton, Donald E.
: APPLICANT: Lipsky, Brian P.

TITLE OF INVENTION: Cytoplasmic Modulators of Integrin
: TITLE OF INVENTION: Binding/Signaling
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
: STREET: 233 South Wacker, 6300 Sears Tower
: CITY: Chicago
: STATE: Illinois
: COUNTRY: USA
: ZIP: 60606
COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/839,581A
: FILING DATE:
: CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
: NAME: Young J. Suh
: REGISTRATION NUMBER: P-41,337
: REFERENCE/DOCKET NUMBER: 27865/33886
TELECOMMUNICATION INFORMATION:
: TELEPHONE: 312-474-6300
: TELEFAX: 312-474-0448
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
: LENGTH: 1700 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
MOLECULE TYPE: DNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 283..1482
US-08-839-581A-1

Query Match
Best Local Similarity 2.8%; Score 56.2; DB 4; Length 1700;
Matches 64; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

83.1%; Pred. No. 4.7e-07;

Qy 17 ccacgcttggtaccggagtcggtcccactagtaacggccgccagtgtgctggaaagtgtt 76
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1691 CCACGCCTTGGTACCGGAGCTCGGATCCACTAGTAACGGGCGGCCCAGTGTGTGCTCTAAAGTTTAC 1632
Qy 77 caaccagtgattatta 93
|||| |||
Db 1631 CAACAGCGTGTTAATTA 1615

RESULT 6
US-08-393-734-3
: Sequence 3, Application US/08393734
: Patent No. 5652224
: GENERAL INFORMATION:
: APPLICANT: Wilson, James M.
: APPLICANT: Kozarsky, Karen F.
: APPLICANT: Strauss, Jerome F.
: TITLE OF INVENTION: Methods and Compositions for Gene
: TITLE OF INVENTION: Therapy for the Treatment of Defects in Lipoprotein
: TITLE OF INVENTION: Metabolism
: NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
: ADDRESSEE: Howson and Howson
: STREET: Spring House Corporate Cntr., PO Box 457
: CITY: Spring House
: STATE: Pennsylvania
: COUNTRY: USA
: ZIP: 19477
COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/393,734
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: UPNH1254USA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 9552 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA

US-393-734-3

Query Match   2.6%;   Score 52.6; DB 1; Length 9552;
Best Local Similarity   86.6%; Pred. No. 1.4e-05;
Matches   58; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy   4 actataggagaccaacgcttggtaccgaagctcggataccactagtaaacggccgccaagtgt 63
           ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4092 ACCTTTGCGTGGATCAAGCTTGGTACCGAGCTCGGATCCACTAGTAACGGCCGCCAGTGT 4151

Qy   64 gctggaa 70
           |||||||
Db   4152 GCTGGAA 4158

RESULT 7
US-08-454-557C-76/c
Sequence 76, Application US/08454557C
Patent No. 5830670
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
APPLICANT: Wands, Jack R.
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,557C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 76:
SEQUENCE CHARACTERISTICS:
LENGTH: 120 base pairs

TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-76

Query Match   2.6%;   Score 52.2; DB 3; Length 120;
Best Local Similarity   71.1%; Pred. No. 1.7e-06;
Matches   69; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy   4 actataggagaccaacgcttggtaccgaagctcggataccactagtaaacggccgccaagtgt 63
           ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   98 ACTCAAGCTATGCATCAACTTGGTACCGAGCTCGGATCCACTAGTAACGGCTGCCAGTGT 39

Qy   64 gctggaaagtgttcaaccagtgattattttgaca 100
           ||| ||||| |||| ||||| |||||||| | ||
Db   38 GCTCTAAAGTATACAATATATATTGTTATATTGTCTATA 2

RESULT 8
US-08-340-426D-76/c
Sequence 76, Application US/08340426D
Patent No. 5948634
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
APPLICANT: Wands, Jack R.
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,426D
FILING DATE: 14-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 76:
SEQUENCE CHARACTERISTICS:
LENGTH: 120 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-340-426D-76

Query Match   2.6%;   Score 52.2; DB 4; Length 120;
Best Local Similarity   71.1%; Pred. No. 1.7e-06;
Matches   69; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy   4 actataggagaccaacgcttggtaccgaagctcggataccactagtaaacggccgccaagtgt 63
           ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   98 ACTCAAGCTATGCATCAACTTGGTACCGAGCTCGGATCCACTAGTAACGGCTGCCAGTGT 39

Qy   64 gctggaaagtgttcaaccagtgattattttgaca 100
           ||| ||||| |||| ||||| |||||||| | ||
Db   38 GCTCTAAAGTATACAATATATATTGTTATATATTGTCTATA 2

RESULT 9
US-08-450-673C-76/c
Sequence 76, Application US/08450673C
Patent No. 5948888
GENERAL INFORMATION:
  APPLICANT: de la Monte, Suzanne
  APPLICANT: Wands, Jack R.
  TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
  TITLE OF INVENTION: of Alzheimer's Disease
  NUMBER OF SEQUENCES: 121
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
    STREET: 1100 New York Avenue, Suite 600
    CITY: Washington
    STATE: D.C.
    COUNTRY: U.S.A.
    ZIP: 20005-3934
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/450,673C
    FILING DATE: 30-MAY-1995
    CLASSIFICATION: 530
  ATTORNEY/AGENT INFORMATION:
    NAME: Ludwig, Steven R.
    REGISTRATION NUMBER: 36,203
    REFERENCE/DOCKET NUMBER: 0609.3840004
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (202) 371-2600
    TELEFAX: (202) 371-2540
  INFORMATION FOR SEQ ID NO: 76:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 120 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: both
      TOPOLOGY: both

Query Match   2.6%;   Score 52.2;   DB 4;   Length 120;
Best Local Similarity   71.1%;   Pred. No. 1.7e-06;
Matches   69;   Conservative   0;   Mismatches   28;   Indels   0;   Gaps   0;

Qy   4 actataggagaccacgccttggtaccgagctcggaatccactagtaacggccgccagtgt 63
       ||| | | | |||||||||||||||||||||||||||||||| ||||||||
Db   98 ACTCAAGCTATGCATCAACTTGGTACCGAGCTCGGATCCACTAGTAACGGCTGCCAGTGT 39

Qy   64 gctggaaagtgttcaaccagtgattatttattgaca 100
       ||| ||||||| ||| |||| | |||| | ||
Db   38 GCTCTAAAGTATACAAATATATTGTTATATTGTCTATA 2

RESULT 10
PCT-US95-17111A-76/c
Sequence 76, Application PCT/TUS9517111A
GENERAL INFORMATION:
  APPLICANT: de la Monte, Suzanne
  APPLICANT: Wands, Jack R.
  TITLE OF INVENTION: Neural Thread Protein Gene Expression and
  TITLE OF INVENTION: Detection of Alzheimer's Disease
  NUMBER OF SEQUENCES: 121
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
    STREET: 1100 New York Avenue, Suite 600
    CITY: Washington
    STATE: D.C.
    COUNTRY: U.S.A.
    ZIP: 20005-3934
  COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/17111A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,426
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 76:
SEQUENCE CHARACTERISTICS:
LENGTH: 120 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
PCT-US95-17111A-76

Query Match   2.6%;   Score 52.2;   DB 6;   Length 120;
Best Local Similarity   71.1%;   Pred. No. 1.7e-06;
Matches   69;   Conservative   0;   Mismatches   28;   Indels   0;   Gaps   0;

Qy   4 actataggagaccacgccttggtaccgagctcggaatccactagtaacggccgccagtgt 63
       ||| | | | |||||||||||||||||||||||||||||||| ||||||||
Db   98 ACTCAAGCTATGCATCAACTTGGTACCGAGCTCGGATCCACTAGTAACGGCTGCCAGTGT 39

Qy   64 gctggaaagtgttcaaccagtgattatttattgaca 100
       ||| ||||||| ||| |||| | |||| | ||
Db   38 GCTCTAAAGTATACAAATATATTGTTATATTGTCTATA 2

RESULT 11
US-08-040-753-1
Sequence 1, Application US/08040753
Patent No. 5464745
GENERAL INFORMATION:
  APPLICANT: Mierendorf, Robert
  APPLICANT: Garber, Richard
  APPLICANT: No. 5464745y, Robert
  APPLICANT: Hammer, Beth
  TITLE OF INVENTION: Protein Ligand Binding
  TITLE OF INVENTION: Region Mapping System
  NUMBER OF SEQUENCES: 2
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Quarles and Brady
    STREET: 1 South Pinckney St., Box 2113
    CITY: Madison
    STATE: WI
    COUNTRY: USA
    ZIP: 53701-2113
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/040,753
    FILING DATE: 19930331
    CLASSIFICATION: 435
  ATTORNEY/AGENT INFORMATION:
    NAME: Seay, Nicholas J
    REGISTRATION NUMBER: 27386
    REFERENCE/DOCKET NUMBER: 70-399-9001-1
  TELECOMMUNICATION INFORMATION:

TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3984 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: pTOPE-1b(+)
FEATURE:
NAME/KEY: promoter
LOCATION: 57..85
OTHER INFORMATION: /function= "T7 Promoter"
FEATURE:
NAME/KEY: RBS
LOCATION: 134..139
FEATURE:
NAME/KEY: CDS
LOCATION: 140..1092
OTHER INFORMATION: /product="5' end of T7 gene 10
OTHER INFORMATION: fusion protein"/note="This coding region is interrupted
OTHER INFORMATION: during cloning by insertion of putative epitope encoding
OTHER INFORMATION: DNA."

US-08-040-753-1

Query Match
Best Local Similarity 94.4%; Pred. No. 1.1e-05;   Score 52.2; DB 1; Length 3984;
Matches 54; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy   14 gaccagcgcttggtaccgagctccggatccactagtaacggccgccagtgtgctggaa 70
          ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 931 GATTCCAAGCTTGGTACCGAGCTCCGGATCCACTAGTAACGGCCGCCAGTGTGCTGGAA 987

RESULT 12
US-08-832-883-68
Sequence 68, Application US/0883288 3
Patent No. 5807681
GENERAL INFORMATION:
APPLICANT: Giordano, Antonio
APPLICANT: Baldi, Alphonso
TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS AND PROGNOSIS
TITLE OF INVENTION: OF CANCER
NUMBER OF SEQUENCES: 115
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.
STREET: Suite 1800 Two Penn Center Plaza
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/832,883
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 8321-13 US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549

INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 9837 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-832-883-68

Query Match
Best Local Similarity 98.1%; Pred. No. 3.2e-05;   Score 51.4; DB 2; Length 9837;
Matches 52; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy   18 cacgcttggtaccgagctccggatccactagtaacggccgccagtgtgctggaa 70
          | |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7170 CAAGCTTGGTACCGAGCTCCGGATCCACTAGTAACGGCCGCCAGTGTGCTGGAA 7222

RESULT 13
US-08-832-877-68
Sequence 68, Application US/0883287 7
Patent No. 5840506
GENERAL INFORMATION:
APPLICANT: Giordano, Antonio
TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS AND PROGNOSIS OF
TITLE OF INVENTION: CANCER
NUMBER OF SEQUENCES: 116
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.
STREET: Suite 1800 Two Penn Center Plaza
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/832,877
FILING DATE:
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 8321-13 US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549

INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 9837 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-832-877-68

Query Match
Best Local Similarity 98.1%; Pred. No. 3.2e-05;   Score 51.4; DB 3; Length 9837;
Matches 52; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy   18 cacgcttggtaccgagctccggatccactagtaacggccgccagtgtgctggaa 70
          | |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7170 CAAGCTTGGTACCGAGCTCCGGATCCACTAGTAACGGCCGCCAGTGTGCTGGAA 7222

RESULT 14
US-08-252-493C-4
Sequence 4, Application US/08252493C

Patent No. 5891645
GENERAL INFORMATION:
; APPLICANT: Rollins, Scott
; APPLICANT: Rother, Russell P.
; APPLICANT: Evans, Mark J.
; APPLICANT: Matis, Louis A.
; TITLE OF INVENTION: PORCINE E-SELECTIN
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seth A. Fidel
; STREET: 25 Science Park, Box 15
; CITY: New Haven
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06511
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 750 Kb storage
; COMPUTER: PC compatible
; OPERATING SYSTEM: DOS 6.2
; SOFTWARE: WordPerfect 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/252,493C
; FILING DATE: June 1, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fidel, Seth A.
; REGISTRATION NUMBER: 38,449
; REFERENCE/DOCKET NUMBER: ALX-138
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203) 776-1790
; TELEFAX: (203) 772-3655
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4059 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Double
; TOPOLOGY: Circular
; MOLECULE TYPE: Other nucleic acid
; DESCRIPTION: Apex-1 Eukaryotic
; DESCRIPTION: Expression Vector
US-08-252-493C-4

● Query Match
; Local Similarity   2.5%;   Score 50;   DB 3;   Length 4059;
Ches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   21 gcttggtaccgagctcggatccactagtaacggccgccagtgtgctggaa 70
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1027 GCTTGGTACCGAGCTCGGATCCACTAGTAACGGCCGCCCAGTGTGCTGGAA 1076

RESULT 15
PCT-US95-07554-4
Sequence 4, Application PC/TUS9507554
; GENERAL INFORMATION:
; APPLICANT: Sandrin, Mauro S.
; APPLICANT: Fodor, William L.
; APPLICANT: Rother, Russell P.
; APPLICANT: Squinto, Stephen P.
; APPLICANT: McKenzie, Ian F. C.
; TITLE OF INVENTION: Methods for Reducing
; TITLE OF INVENTION: Hyperacute Rejection of Xenografts
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maurice M. Klee
; STREET: 1951 Burr Street
; CITY: Fairfield
; STATE: Connecticut
; COUNTRY: USA

; ZIP: 06430
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 750 Kb storage
; COMPUTER: Dell 486/50
; OPERATING SYSTEM: DOS 6.2
; SOFTWARE: WordPerfect 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07554
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/260,201
; FILING DATE: June 15, 1994
; CLASSIFICATION:
; APPLICATION NUMBER: 08/278,282
; FILING DATE: July 21, 1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Klee, Maurice M.
; REGISTRATION NUMBER: 30,399
; REFERENCE/DOCKET NUMBER: ALX-144.1PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203) 255-1400
; TELEFAX: (203) 254-1101
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4059 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Double
; TOPOLOGY: Circular
; MOLECULE TYPE: Other nucleic acid
; DESCRIPTION: Apex-1 Eukaryotic
; DESCRIPTION:
; ANTI-SENSE: No
; PCT-US95-07554-4

Query Match
Best Local Similarity   2.5%;   Score 50;   DB 6;   Length 4059;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   21 gcttggtaccgagctcggatccactagtaacggccgccagtgtgctggaa 70
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1027 GCTTGGTACCGAGCTCGGATCCACTAGTAACGGCCGCCCAGTGTGCTGGAA 1076

Search completed: March 24, 2000, 13:20:42
Job time: 8033 sec

```
RESULT   1
US-08-374-483-1
; Sequence 1, Application US/08374483
; Patent No. 5880102
;  GENERAL INFORMATION:
;    APPLICANT:  GEORGE, SAMUEL E.
;    APPLICANT:  BLAZING, MICHAEL A.
;    TITLE OF INVENTION:  ADENOVIRAL VECTOR SYSTEM
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIXON & VANDERHYE P.C.
;      STREET:  1100 NORTH GLEBE ROAD, 8TH FLOOR
;      CITY:  ARLINGTON
;      STATE:  VIRGINIA
;      COUNTRY:  U.S.A.
;      ZIP:  22201-4714
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/374,483
;      FILING DATE:  17-JAN-1995
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  WILSON, MARY J.
;      REGISTRATION NUMBER:  32,955
;      REFERENCE/DOCKET NUMBER:  1579-83
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 816-4000
;      TELEFAX:  (703) 816-4100
;      TELEX:  200797 NIXN UR
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6420 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-374-483-1
```

```
    Query Match              3.5%;  Score 69.6;  DB 3;  Length 6420;
    Best Local Similarity   77.8%;  Pred. No. 1.1e-10;
    Matches    84;  Conservative    0;  Mismatches    24;  Indels    0;  Gaps    0;

Qy      1 ctcactatagggagacccacgcttggtaccgagctcggatccactagtaacggccgccag 60
          |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db    975 CTCACTATAGGGAGACCCAAGCTTGGTACCGAGCTCGGATCCACTAGTAACGGCCGCCAG 1034
```

'89-38_

```
Qy     61 tgtgctggaaagtgttcaaccagtgattatttatttgacagcaactgg 108
          ||||||||||   ||| |   |   ||     || | || ||| |
Db   1035 TGTGCTGGAATTAATTCTGCAGATATCCATCACACTGGCGGCCGCTCG 1082
```

THIS PAGE BLANK (USPTO)

Patent No. 5464745
GENERAL INFORMATION:
APPLICANT: Mierendorf, Robert
APPLICANT: Garber, Richard
APPLICANT: NO. 5464745y, Robert
APPLICANT: Hammer, Robert
APPLICANT: Hammer, Beth
TITLE OF INVENTION: Protein Ligand Binding
TITLE OF INVENTION: Region Mapping System
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles and Brady
STREET: 1 South Pinckney St., Box 2113
CITY: Madison
STATE: WI
COUNTRY: USA
53701-2113
C...ER READABLE FORM:
...IUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/040,753
FILING DATE: 1930331
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 70-399-9001-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3984 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: circular
...ECULE TYPE: DNA (genomic)
...HETICAL: NO

ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: pTOPE-1b(+)
FEATURE:
NAME/KEY: promoter
LOCATION: 67..85
OTHER INFORMATION: /function= "T7 Promoter"
FEATURE:
NAME/KEY: RBS
LOCATION: 134..139
FEATURE:
NAME/KEY: CDS
LOCATION: 148..1092
OTHER INFORMATION: /product= "5' end of T7 gene 10
OTHER INFORMATION: fusion protein"/note= "This coding region is interrupted
OTHER INFORMATION: during cloning by insertion of putative epitope encoding
OTHER INFORMATION: DNA.
"
US-08-040-753-1

Query Match   2.0%;   Score 50;   DB 1;   Length 3984;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches   50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 gcttggttaccggagtctcggatcccactagtcaacggccgccaagtgtgctggaa 50
        ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db 938 GCTTGGTTACCGGAGCTCGGATCCACTAGTTAACGGCCGCCGCCAGTGTGCTGGAA 987

RESULT 14
US-08-252-493C-⌐
Sequen⌐

.CTRA 7232

```
US        CVU89672    5726 bp    DNA             SYN       21-MAR-1997
INITION   Cloning vector pIRES1hyg, complete plasmid sequence.
ESSION    U89672
SION      U89672.1  GI:1899166
WORDS     .
RCE       bicistronic.
RGANISM   unidentified cloning vector.
          unidentified cloning vector
          artificial sequence; vectors.
ERENCE    1  (bases 1 to 5726)
THORS     Rees,S., Coote,J., Stables,J., Goodson,S., Harris,S. and Lee,M.G.
ITLE      Bicistronic vector for the creation of stable mammalian cell lines
          that predisposes all antibiotic-resistant cells to express
          recombinant protein
JRNAL     BioTechniques 20 (1), 102-104 (1996)
RENCE     2  (bases 1 to 5726)
THORS     Lou,Y. and Holtz,A.
TLE       Direct Submission
JRNAL     Submitted (14-FEB-1997) CLONTECH Laboratories, Inc., 1020 East
          Meadow Circle, Palo Alto, CA 94303-4230, USA
IENT      This vector can be obtained from CLONTECH Laboratories, Inc., 1020
          East Meadow Circle, Palo Alto, CA 94303, USA. To place an order
          call (415) 424-8222 or (800) 662-2566, extension 1. International
          customers, please contact your local distributor. For technical
          information, call (415) 424- 8222 or (800) 662-2566, extension 3.
          This sequence has been compiled from information in the sequence
          databases, published literature and other sources, together with
          partial sequences obtained by CLONTECH. If you suspect there is an
          error in this sequence, please contact CLONTECH's Technical Service
          Department at (415) 424-8222 or (800) 662-2566, extension 3, or
          E-mail TECH@CLONTECH.COM.
URES      Location/Qualifiers
source    1..5726
          /organism="unidentified cloning vector"
          /db_xref="taxon:45196"
Intron    938..1233
          /note="synthetic intron"
misc_feature  1273..1858
          /note="internal ribosome entry site"
CDS       1872..2906
          /codon_start=1
          /transl_table=11
          /product="hygromycin B phosphotransferase"
          /protein_id="AAB49979.1"
          /db_xref="GI:1899167"
          /translation="MDRSGKPELTATSVEKFLIEKFDSVSDLMQLSEGEESRAFSFDV
          GGRGYVLRVNSCADGFYKDRYVYRHFASAALPIPEVLDIGEFSESLTYCISRRAQGVT
          LQDLPETELPAVLQPVAEANDAIAAADLSQTSGFGPGPGQIGOYTTWRDFICAIADP
          HVYHNGTVMDDTVSASVAQALDELMLWAEDCPEVRHLVHADFGSNNVLTDNGRITAVI
          DWSEAMFGDSQYEVANIFFWRPWLACMEQQTRYFERRHPELAGSPRLRAYMLRIGLDQ
          LYQSLVDGNFDDAAWAQGRCDAIVRSGAGTVGRIQIARRSAAVWTDGCVEVIADSGNR
          RPSTRPRAKE"
gene      4730..5591
          /gene="bla"
CDS       complement(4730..5590)
          /gene="bla"
          /EC_number="3.5.2.6"
          /function="ampicillin-resistance"
          /codon_start=1
          /transl_table=11
          /product="beta-lactamase"
          /protein_id="AAB49980.1"
          /db_xref="GI:1899168"
          /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
          IELDLNSGKILESFRPEERFPMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
          YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTIGGPKELTAFLHNGDHVTRL
          DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
          LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
          EIGASLIKHW"
COUNT     1328 a    1482 c    1518 g    1398 t
```

Query Match
Best Local Similarity   3.4%;   Score 69;   DB 14;   Length 5726;
Matches   72;   Conservative   0;   Pred. No. 3. 8e-06;
                                   0;   Mismatches   5;   Indels   0;   Gaps   0;

```
Qy     1 ctcaactataggggagaccacgcttggtaccagagcctggatccacactagtaacggccgccag 60
          ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db   871 CTCAACTATAGGGAGACCCCAAGCTTGGTACCAGAGCTCGGATCCACACTAGTAACGGCCGCCAG 930

Qy    61 tgtgctgggaaagtgttc 77
          |||||||||| ||  |||
Db   931 TGTGCTGGAATTAATTC 947
```

us-09-183-789-38.rge

OM nucleic - nucleic search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

Run on: March 24, 2000, 04:02:39 ; Search time 5911.47 Seconds
                       (without alignments)
                       -1038.072 Million cell updates/sec

Title: US-09-183-789-38
Perfect score: 2021
Sequence: 1 ctcactataggagacccac.............aaaaaaaaaaaaaaaaaaa 2021
Scoring table: IDENTITY_NUC
Searched: 821193 seqs, -1518192014 residues
Searched: GenEmbl:*
abase :   GenEmbl:*
Word size :   0

Number of hits that pass the threshold :   1642386

50: gb_pl3:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	209.6	10.4	226280	41	AC011366	AC011366 Homo sapi
2	209.6	10.4	215387	42	AC008417	AC008417 Homo sapi
3	119.2	5.9	476	13	HS222P2T	AL032554 H.sapiens
4	86.4	4.3	1738	10	HSM801162	AL117626 Homo sapi
5	86.4	4.3	3690	40	AF097645	AF097645 Homo sapi
6	84.8	4.0	2110	40	AF141326	AF141326 Homo sapi
7	80	4.0	6346	12	MMU57368	U57368 Mus musculu
8	69	3.4	5726	14	CVU89672	U89672 Cloning vec
9	68.4	3.4	6802	5	A49700	A49700 Sequence 6
10	68.4	3.4	10443	5	A58521	A58521 Sequence 25
11	68.4	3.4	3840	14	EVE132038	AJ132038 Expressio
12	68.4	3.4	3986	14	PCDNA32EO	X90639 Cloning vec
13	68.4	3.4	7290	14	XXU43958	U43958 Cloning vec
14	66.8	3.3	4525	5	AR062871	AR062871 Sequence
15	62.4	3.1	161	10	I66494	I66494 Sequence 14
16	60.4	3.0	7218	5	HSU31738	U31738 Homo sapien
17	59.6	2.9	1123	43	I65436	I65436 Sequence 3
18	59.2	2.9	184553	43	AC013452	AC013452 Homo sapi
19	55.8	2.8	510	13	G12382	G12382 UNH231 Tila
20	54.8	2.7	202342	41	AC009473	AC009473 Homo sapi
21	54.4	2.7	719	2	AF018054	AF018054 Uncleture
22	53.6	2.7	175153	42	AC010978	AC010978 Homo sapi
23	53.4	2.6	689	2	AF018044	AF018044 Uncleture
24	53.2	2.6	175153	42	AC010978	AC010978 Homo sapi
25	53	2.6	1102	14	AF137070	AF137070 Sphenosty
26	53	2.6	173630	42	AC010902	AC010902 Homo sapi
27	52.6	2.6	9592	13	I58669	I58669 Sequence 3
28	52.6	2.6	357	13	G31347	G31347 UNH-009 Til
29	52.6	2.6	162575	33	AC004086	AC004086 Homo sapi
30	52.6	2.6	52359	41	AC010772	AC010772 Homo sapi
31	52.4	2.6	719	2	AF018052	AF018052 Uncleture
32	52.4	2.6	731	2	AF114621	AF114621 Uncleture
33	52.4	2.6	339	2	AF114621	AF114621 Uncleture
34	52.4	2.6	1566	8	AF131673	AF131673 Zea mays
35	52.4	2.6	1623	8	AF131683	AF131683 Zea mays
36	52.4	2.6	3418	9	HSP58NKRC	AJ000542 Homo sapi
37	52.4	2.6	1577	10	AB029348	AB029348 Homo sapi
38	52.4	2.6	4224	12	AF033115	AF033115 Mus muscu
39	52.4	2.6	1673	12	D88984	D88984 Mus musculu
40	52.4	2.6	3942	14	AF073995	AF073995 Vibrio ch
41	52.4	2.6	1199	16	DVU93295	U93295 Dengue viru
42	52.4	2.6	1197	16	DVU93301	U93301 Dengue viru
43	52.4	2.6	1200	16	DVU93302	U93302 Dengue viru
44	52.2	2.6	120	5	AR051506	AR051506 Sequence
45	52.2	2.6	3984	5	I15353	I15353 Sequence 1

ALIGNMENTS

RESULT   1
AC011366
LOCUS   AC011366   226280 bp   DNA   HTG
DEFINITION Homo sapiens chromosome 5 clone CIT-HSPC_568L21, *** SEQUENCING IN
            PROGRESS ***, 82 unordered pieces.
ACCESSION   AC011366
VERSION   AC011366.1 GI:6013575
KEYWORDS   HTG; HTGS_PHASE1.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.

gb_ba1:*		1:			06-OCT-1999
gb_ba2:*	2:				
gb_om:*	3:				
gb_ov:*	4:				
gb_pat:*	5:				
gb_ph:*	6:				
gb_pl1:*	7:				
gb_pl2:*	8:				
gb_pr1:*	9:				
gb_pr2:*	10:				
gb_pr3:*	11:				
gb_ro:*	12:				
gb_sts:*	13:				
gb_sy:*	14:				
gb_un:*	15:				
gb_vi:*	16:				
em_fun:*	17:				
em_hum1:*	18:				
em_hum2:*	19:				
em_in:*	20:				
em_om:*	21:				
em_or:*	22:				
em_ov:*	23:				
em_pat:*	24:				
em_ph:*	25:				
em_pl:*	26:				
em_ro:*	27:				
em_sts:*	28:				
em_sy:*	29:				
em_un:*	30:				
em_vi:*	31:				
gb_htg1:*	32:				
gb_htg2:*	33:				
gb_in1:*	34:				
gb_in2:*	35:				
gb_ba1:*	36:				
gb_ba2:*	37:				
em_ba1:*	38:				
em_ba2:*	39:				
em_hum3:*	40:				
gb_pr4:*	41:				
gb_htg3:*	42:				
gb_htg4:*	43:				
gb_htg5:*	44:				
gb_htg6:*	45:				
gb_htg7:*	46:				
em_htg1:*	47:				
em_htg2:*	48:				
em_htg3:*	49:				
em_hum5:*					

```
REFERENCE   1  (bases 1 to 226280)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 5
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 226280)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-OCT-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     www.jgi.doe.gov.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 82 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the contigs are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1      645: contig of 645 bp in length
            *      646     1347: gap of unknown length
            *     1348     2001: contig of 654 bp in length
            *     2002     2675: gap of unknown length
            *     2676     3336: contig of 661 bp in length
            *     3337     3966: gap of unknown length
            *     3967     4625: contig of 659 bp in length
            *     4626     5279: gap of unknown length
            *     5280     5967: contig of 688 bp in length
            *     5968     6664: gap of unknown length
            *     6665     7351: contig of 687 bp in length
            *     7352     8034: gap of unknown length
            *     8035     8717: contig of 683 bp in length
            *     8718     9374: gap of unknown length
            *     9375    10049: contig of 675 bp in length
            *    10050    10721: gap of unknown length
            *    10722    11432: contig of 711 bp in length
            *    11433    12070: gap of unknown length
            *    12071    12682: contig of 612 bp in length
            *    12683    13282: gap of unknown length
            *    13283    13605: contig of 323 bp in length
            *    13606    13659: gap of unknown length
            *    13660    14327: contig of 668 bp in length
            *    14328    14983: gap of unknown length
            *    14984    15700: contig of 717 bp in length
            *    15701    16412: gap of unknown length
            *    16413    16677: contig of 265 bp in length
            *    16678    17340: gap of unknown length
            *    17341    18050: contig of 710 bp in length
            *    18051    18724: gap of unknown length
            *    18725    19449: contig of 725 bp in length
            *    19450    19624: gap of unknown length
            *    19625    20309: contig of 685 bp in length
            *    20310    20938: gap of unknown length
            *    20939    21559: contig of 621 bp in length
            *    21560    22223: gap of unknown length
            *    22224    22950: contig of 727 bp in length
            *    22951    23584: gap of unknown length
            *    23585    24260: contig of 676 bp in length
            *    24261    24888: gap of unknown length
            *    24889    25558: contig of 670 bp in length
            *    25559    26234: gap of unknown length
            *    26235    26911: contig of 677 bp in length
            *    26912    27669: gap of unknown length
            *    27670    28345: contig of 676 bp in length
            *    28346    28963: gap of unknown length
            *    28964    29571: contig of 608 bp in length
            *    29572    30211: gap of unknown length
            *    30212    30303: contig of 92 bp in length
            *    30304    30963: gap of unknown length
            *    30964    31660: contig of 697 bp in length
            *    31661    32323: gap of unknown length
            *    32324    33002: contig of 679 bp in length
            *    33003    33685: gap of unknown length
            *    33686    34355: contig of 670 bp in length
            *    34356    35026: gap of unknown length
            *    35027    35708: contig of 682 bp in length
            *    35709    35802: gap of unknown length
            *    35803    36426: contig of 624 bp in length
            *    36427    37156: gap of unknown length
            *    37157    37912: contig of 756 bp in length
            *    37913    38580: gap of unknown length
            *    38581    39278: contig of 698 bp in length
            *    39279    39943: gap of unknown length
```

FEATURES   Location/Qualifiers
    source   1..226280
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="5"
                /clone="CIT-HSPC_568L21"
    *   100124 226280: contig of 126157 bp in length.
    *   73973   100123: contig of 26151 bp in length
    *   47850   73972: contig of 26123 bp in length
    *   47488   47849: contig of 362 bp in length
    *   46599   47487: contig of 889 bp in length
    *   46468   46598: contig of 131 bp in length
    *   45731   46467: contig of 737 bp in length
    *   45353   45730: contig of 378 bp in length
    *   44771   45352: contig of 582 bp in length
    *   43654   44770: contig of 1117 bp in length
    *   43551   43653: contig of 103 bp in length
    *   43454   43550: contig of 97 bp in length
    *   43453: contig of 1122 bp in length
    *   42332   43452: contig of 88 bp in length
    *   42244   42331: contig of 398 bp in length
    *   41846   42243: contig of 1052 bp in length
    *   40794   41845: contig of 666 bp in length
    *   40128   40793: contig of 184 bp in length
    *   39944   40127: contig of 184 bp in length

ORIGIN
BASE COUNT   69691 a   46147 c   45719 g   64623 t   100 others

Query Match   10.4%;   Score 209.6;   DB 41;   Length 226280;
Best Local Similarity 63.2%;   Pred. No. 7.7e-39;
Matches 373; Conservative 0; Mismatches 184; Indels 33; Gaps 2;

Qy 1366 tatgacagaacattgtatgcccaccaatggattctttcctcacgacttcacaag 1425
          |||||||||| ||||| ||| |||||| |||||||| |||||||||| |||||
Db 207374 TATGGCAGGAGGATGGTATTCCACCAAGCCAATTGGATTCTCGGTTTGATGACTTCAGTGG 207433

Qy 1426 tctcagcaaagatgagctgcttacaaacctgatagtaatgaattgcggtaggcaccaa 1485
          |||||| || | ||| ||| |||||||||| | | | | ||||||| ||||
Db 207434 TTGCACAAAAGAATGGGCGGCTGATCCAGAGAAACCTGTAGAAAATGCACCTGTAGGAGGAAATCAT 207493

Qy 1486 aaactaccagtgtctctgcaggtgacccaccacagttacagtataatgtcttcggtgaaactgt 1545
          ||| || || |||||||| | | ||||||| || ||||||||
Db 207494 TACCACGCAATTTCTCTGGAGATGACCTAAAAGTCACAGAAAT 207536

Qy 1546 gccaatacacacaaatatctcctgccatgcaaaaataattaatgatgatataaaata 1605
          ||||||||| | ||| || || ||||| |||||||||| ||||||||||||
Db 207536 ACTCCCTTTTCAAAATGTCAAGAAGAGAAATATTAATGCTGATATTAAAATG 207583

Qy 1606 tcaattaatgaaagaagttcgaaggttcgggcaaaatcattgaaagaagaatttcattttgct 1665
          ||||||| || ||||| ||| | |||||| ||||||| ||| |||||||||||||||||
Db 207584 TCAATTAGTTGAGGAAAATTCACACTTTAGACGAAAATATGAAAAACTCTTCAAATTGCT 207643

Qy 1666 tgaagaggtacaagatctatgaaagtcaagagacaattgttgaatttaccatcaagga 1725
          ||||| ||| ||||| |||||| ||| |||||||||| ||||| | |||||||||||||
Db 207644 TGAAGGACTGCAAGGACCTATAGAGTCGAAAGTCAAGAAATGAAAATGATTTTTGAATCCATCATCAAGGA 207703

RESULT 2

LOCUS   AC008417   215387 bp   DNA   HTG   31-OCT-1999
DEFINITION Homo sapiens chromosome 5 clone CIT-HSPC_28708, *** SEQUENCING IN
            PROGRESS ***, 5 ordered pieces.
ACCESSION   AC008417
VERSION   AC008417.2 GI:6165190
KEYWORDS   HTG; HTGS_PHASE2.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 215387)
  AUTHORS   DOE Joint Genome Institute.
  TITLE   Sequencing of Human Chromosome 5
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 215387)
  AUTHORS   DOE Joint Genome Institute.
  TITLE   Direct Submission
  JOURNAL   Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT   On Oct 31, 1999 this sequence version replaced gi:5686649.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 5 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submitter.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *   1 10088: contig of 10088 bp in length
            *   gap of unknown length
            *   10989 107771: contig of 96783 bp in length
            *   gap of unknown length
            *   107772 193020: contig of 85249 bp in length
            *   gap of unknown length
            *   193021 195835: contig of 2815 bp in length
            *   gap of unknown length
            *   195836 215387: contig of 19552 bp in length.
FEATURES   Location/Qualifiers
    source   1..215387
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="5"
                /clone="CIT-HSPC_28708"
ORIGIN
BASE COUNT   74310 a   41080 c   38337 g   61386 t   274 others

Query Match   10.4%;   Score 209.6;   DB 42;   Length 215387;
Best Local Similarity 63.2%;   Pred. No. 7.7e-39;
Matches 373; Conservative 0; Mismatches 184; Indels 33; Gaps 2;

Qy 1726 agcagcaagtgttaaaaagttgtcttaattcagcaactcgagaaggcggcttaaagaat 1785
          ||| || |||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||||||||
Db 207704 AGCCACAGATTATTGAGACGAGACGAGACTTAATTCACCACCTTGAGAAGAAAACTGGAACAAAT 207763

Qy 1786 agattcccactgccatccagaaaagttaagcacacatgagaaaagatattgttagtg 1845
          ||||||||| | || ||||||||||| ||||| | ||||||| | |||||| |||| |
Db 207764 GATTTCTGGCCTACTTGTTCAAGAAGGATCATCACCGCGCCGAAATGCCATAATCTTGTTAATG 207823

Qy 1846 ---caaagaccaaggagaaacaagacatatgtcgtaggatggaacaggttattgctgaa 1902
             |||| | || |||| || |||| | |||||||||| ||||| ||||||| |||||||
Db 207824 ACTGAGGAGGAGGAAAAAGAGATCAAAATTGCTGTTTTCTAGAATGGAGCAGGATATTGCTGAA 207883

Qy 1903 gctccctataatcctgaaatgaagagaattccctccagaagctacgaa 1952
          ||||| | ||| ||| |||||||||||| ||||||||||| || |||
Db 207884 GCCTCCTAGCACATGTGTTAGTGAATAAAAAATGGCCTTCCAGAGGCTAAGACA 207933

Qy 1366 tatgacagaacattgtatgcccaccaatggattctttcctcacgacttcacaag 1425
          ||||||||| ||||| ||| |||||| ||||||||||||||||| ||||||||
Db   ...

```
Db   24631  TATGGCAGGAGATGGTATTCCACCAAGCCAATTGGATTCTCGGTTGATGAGACTTCAGTGG  24690

Qy    1426  tctcagcaaagatgagctgcttacacaaacctgataagtaatgaattgcggtaggcaccaa  1485
            ||||||||||||| ||||| || ||  |||||| |||| ||  ||| |||  |||||| ||
Db   24691  TTGCAGCAAAGATGGGCTGATGCAGAAAACCTGGTATGAGAAATGCACCTGTAGGAGGAATCAT  24750

Qy    1486  aaactacagtgtctctgcagtgaccaccagttacagtaatgtcttcggtggaaactgt  1545
            |||| || |||||||||  || ||  ||||| |||||| ||||| |||||
Db   24751  TACCAGCAATTCTCTGGAGATGACCTAAAAGTCACAGAAAGTCACAGAAAT        24793

Qy    1546  gccaaatacacacaaatatctcctgccatggcaaaaaaaaaattaatgatgatataaata  1605
            || |   |||| ||  ||||||||  |||   |||||||||| |||| |||||||||||||
Db   24793  ------------ACTCCCTTTCCAAAAATGTCAAGAGAAAATTAATCTGATATAAAATG   24840

Qy    1606  tcaattaattaatgaaaagttcgaaagtttgggcaaaaatttcattttgt  1665
            ||||||| | |||||||| || |||| |||||||||||| ||||| |||||
Db   24841  TCAATTAGTGAAAGAAAATTCAACACTTAGCAGAAATATGAAAAACTCTTCAAATTGCT   24900

Qy    1666  tgaagagggtacaagaggatctatgaaagtcaaagagagacaaatttgttgaattttaccatcaagga  1725
            |||||||  | |||| | ||| ||  ||||||||| || ||||| ||||||||||  |||||||||||
Db   24901  TGAAGAGACTCCAAGGACCTATAGAGTCAAGAAATAGATTTTTTGAATCCATCATCAAGGA  24960

Qy    1726  agcagcaaggtttaaaaaagttgtcttaattcatcagaaaggcagggcgcttaagagaaat  1785
            |||| ||||| ||||| |||||| ||| ||||| |||||| || | | | |||||| ||||
Db   24961  AGCCACAAGATTTATTGAGACAGCAGAACGAGACTTAATTCACGAAGAGGGAACAAAT   25020

Qy    1786  agattcccacactgccatctcagaaaagttaagcacatgagaaaagataataattgtggttagtg  1845
            || ||| | | |||| ||||||||||  ||||| |||| ||||| || ||| ||||  | ||||||
Db   25021  GAATTCGTGGCTACTTGTTCAAGAAGGATCATCTCACCGACAAATGCATAATCTTGTTAATG  25080

Qy    1846  ---caaagaccaaggagaaacaaggacatatgtctgtaggatggagacaggtgttattgctgaa  1902
               ||||| ||||| | ||||| ||| || | |  || ||| |||||| ||| | ||| |||||
Db   25081  ACTGAGGAGGACAAAAAGGACCAAAATTGCTGTTTTCTTAGAATGGGACAGGATATATTCTGAA  25140

Qy    1903  gctccctatataatcctgaaatgaaagaaagaatcccctccagaagctacgaaa  1952
            ||||| ||| ||| | || ||||| | || | ||||||||||||||| | |||
Db   25141  GCCTCCTAGCATGTGTTAGTGAATTAAAATGGCCTTCCAGAGGCTAAGACA   25190
```

```
Qy    1512  ca  1513
            ||
Db     475  TA  476
```

RESULT 3
LOCUS   HS228P2T   476 bp   DNA   STS   29-OCT-1998
DEFINITION H.sapiens STS from genomic clone 228P2, sequence tagged site.
ACCESSION   AL032554
VERSION   AL032554.1 GI:3810765
KEYWORDS   STS; single read.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 476)
  AUTHORS   Ross,M., Dunham,A., Huckle,E., Taylor,R. and Hunt,S.
  TITLE   Direct Submission
  JOURNAL   Submitted (28-OCT-1998) E-mail contact: humquery@sanger.ac.uk
            Marker stSG55154 (Primer A : TGGGATTTCTCCTGGTGTA; Primer B :
            CACTTTCTTTCCCCTCCACA; amplimer size : 175 bp) is from sequence
            generated from the T7 end of PAC 228P2. 228P2 is part of the
            bacterial clone contigs constructed by the Chromosome X Mapping
            Group.(http://www.sanger.ac.uk/HGP/ChrX/) 228P2 is from the library
            constructed at the Roswell Park Cancer Institute by the group of
            Pieter de Jong.
            For further details see http://bacpac.med.buffalo.edu/.
FEATURES   Location/Qualifiers
     source   1..476
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /clone="228P2"
BASE COUNT   129 a   99 c   109 g   139 t
ORIGIN

Query Match   5.9%;   Score 119.2;   DB 13;   Length 476;

Best Local Similarity 173; Conservative 71.5%; Pred. No. 8.2e-18;
Matches   173; Conservative 71.5%; 0; Mismatches 63; Indels 6; Gaps 1;

```
Qy    1272  ctgtcattcacgatatccaggagaggagggaaaatgatcaaacccctcctgatggct  1331
            ||||||||||| || |||||||||| || | |  ||| | ||| ||||||||||| ||
Db     241  CTGTCATTCATGAGAGGTCTATGTGGACGGGGAAGAAAGTGTTCAAACACCAACTGATGGCT  300

Qy    1332  tcctgtcaaattctgattccacggagctgataaaatatgaacagttgtcaccca  1391
            ||||| || |||||| || |||  |||| | || ||| ||| | ||| | ||||
Db     301  TTTACTCAAAATCGTGTTTCACCAGCCAGCT----TATGACAGGACATGCTATTCCACCCA  354

Qy    1392  atgcattggtttcttctctcacgacttcacgaagtctcagcaaagagtgagctgcttaca  1451
            ||| |||||| |||||||||||||||||||| | ||||||| ||| | | |||| |||||
Db     355  GCCAATTGGATTCTGCAGTGATGACTTCACTGGTTTCAGCAAAGATATGGATGCAGA  414

Qy    1452  aacctgataagtaatgattttgcggtaggtcacccaaaactacagtgtctctgcaggtgacc  1511
            ||||| |||| ||| ||||| |||| ||| |||||||||  |||| ||||| || | ||||
Db     415  AACCTGTAGGAGGAATGCACCTGTGGGGGAAACGTTACCAGCAGTTTCTCTGGAGATGACC  474
```

```
Qy    1512  ca  1513
Db     475  TA  476
```

RESULT 4
LOCUS   HSM801162   1738 bp   mRNA   PRI   15-SEP-1999
DEFINITION Homo sapiens mRNA; cDNA DKFZp434B105 (from clone DKFZp434B105).
ACCESSION   AL117626
VERSION   AL117626.1 GI:5912207
KEYWORDS   .
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1738)
  AUTHORS   Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and
            Wiemann,S.
  TITLE   Direct Submission
  JOURNAL   Submitted (15-SEP-1999) MIPS, Am Klopferspitz 18a D-82152
            Martinsried, GERMANY
COMMENT   Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); sequenced by Qiagen within the cDNA
            sequencing consortium of the German Genome Project. This clone is
            available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES   Location/Qualifiers
     source   1..1738
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="434 (synonym: htes3). Vector pSport1; host
                     DH10B: sites NotI + SalI"
                     /clone="DKFZp434B105"
                     /dev_stage="adult"
                     /tissue_type="testis"
     gene   1..1433
                     /gene="DKFZp434B105"
     CDS   <1..1433
                     /gene="DKFZp434B105"
                     /note="strong similarity to C-terminus of mouse EGF repeat
                     transmembrane protein"
                     /codon_start=3
                     /product="hypothetical protein"
                     /protein_id="CAB56020.1"
                     /db_xref="GI:5912208"
                     /translation="ASGILKTMPPYLGPLKKAVRMGAPNLIADSMEYGLSISVISY
                     LKKLSQQAKIESDRVIGSVGKKVVQEETGIKVRSRSHGLSMAYRKDFDQQLLQGISEDVP
                     HRLLDLNMKEYTGTOVALLNKDLKPOTFRNAYDIPRNLLDHLTRMRSNLLKSTRFEL
                     KGODEDOVHSVPIAQMGNYOEYLKQVPSPIRELEDOPRRLHTTGNPEFKLDKKGMMID
                     EADEFVAGPOQNKHKRPGEPENMOGIPKRRCMKSPLLRGROOMPAVVNNHIGGKGPPATT
                     QAOPDLIKPLPLHKISETINDSIHDVENHVAPOLSSDITPNAMDTEFSASSPASLL
                     ERPTNHMEALGHDHLHLGTNDLTVGGFLENHEEPRDKEQCAEENIPASSLNKGKKLMHCR
```

SHEEVNTELKAQIMKEIRKPGRKYERIFTLLKHVQGSLQTRLIFLQNVIKEASRFKKR
MIIEQLENFLDEIHRRANQINIHNSN"

ORIGIN
BASE COUNT   660 a   325 c   317 g   436 t
1679

polyA_site

Query Match   4.3%;   Score 86.4;   DB 10;   Length 1738;
Best Local Similarity   58.6%;   Pred. No. 3.5e-10;
Matches 150;   Conservative 0;   Mismatches 106;   Indels 0;   Gaps 0;

Qy 1581 aaaaaattaatgatgataaaatcaattaatgaaagaaagttcgaaagtttgggcaa 1640
        || |||||||||||||||||||||||||||||| | |||||||| | |||||||| ||
Db 1186 AAGAGGTCAATACTGAACTAAAAGCACAAATAATGAAAGAGATCCGAAAGCCAGGAAGAA 1245

Qy 1641 attatgaaagaattttcattttgcttgaagaaggtacaaggatctcatgaaagtcaagagac 1700
        | |||||||| |||| ||| ||||| |||||| || || |||||| |||||||| ||||||
Db 1246 AATATGAAAGAGAATCTTCACTTTACTGAAGCATGTGCAAGGCAGTTTACAACAAGACTAA 1305

Qy 1701 aatttgttgaatttacctcaaggaagcaggtttaaaaaaagttgtcttaattcgc 1760
        ||| || || |||| || || ||||||||||||||||||| ||| ||||||||||
Db 1306 TATTTTTACAAAATGTCATTAAAGAAGCATCAAGGTTTAAAAAAACGAATGCTAATAGAAC 1365

Qy 1761 aactcgagaaggcgcttaaagaaatagattcccactgccatctcaagaaaagttaagcaca 1820
        ||||| ||||| | ||| || |||||| |||| | | |||||||||||||||| |||||||
Db 1366 AACTGGAGAACTTCTTGGATGAAAATTCATCGAGAGCCAATCAGATCAACCATATTAATA 1425

Qy 1821 tgagaaaagatatt 1836
        ||| ||||| || |
Db 1426 GCAATTAAAAGAAAAT 1441

RESULT 5
LOCUS   AF097645   3690 bp   mRNA   PRI   19-OCT-1999
DEFINITION Homo sapiens candidate tumor suppressor protein DICE1 mRNA,
           complete cds.
ACCESSION AF097645
VERSION   AF097645.1   GI:6062873
KEYWORDS   .
SOURCE   human.
ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
           Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3690)
AUTHORS   Wieland,I., Arden,K.C., Michels,D., Klein-Hitpass,L., Bohm,M.,
           Viars,C.S. and Weidle,U.H.
TITLE   Isolation of DICE1: a gene frequently affected by LOH and
           downregulated in lung carcinomas
JOURNAL   Oncogene 18 (32), 4530-4537 (1999)
MEDLINE   9398588
REFERENCE 2 (bases 1 to 3690)
AUTHORS   Wieland,I. and Weidle,U.
TITLE   Direct Submission
JOURNAL   Submitted (06-OCT-1998) Institute for Cell Biology (Cancer
           Research), University of Essen Medical School, Virchowstr. 173,
           Essen, NRW D-45122 Essen, Germany
FEATURES   Location/Qualifiers
     source   1..3690
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /chromosome="13"
                    /map="13q14.12-q14.2"
                    /tissue_type="brain"
                    /dev_stage="fetal"
     CDS   477..3140
                    /note="deleted in lung cancer"
                    /codon_start=1
                    /product="candidate tumor suppressor protein DICE1"
                    /protein_id="AAF03046.1"
                    /db_xref="GI:6062874"
                    /translation="MPILLFLIDTSASMNQRSHLGTTYLDTAKGAVETFMKLRARDPA
                    SRGDRYMLVTFEEPYAIKAGWKENHATFMNELKNLQAELTTLGQSLRTAFDLLNLN

RLVTGIDNYGQGRNPFLEPAIIITITDGSKLTTTSGVQDELHLPLNSPLPGSELTKE
PFRWDQRLFPALVLRLPGTMSVESEOLTGVPLLDDSAITPMCEVTGGRSYSVCSPRMLNQ
CLRSLVQKVQSGVVINFEKAGPDDSPVEDGQPDISRPGSQPWHSCHKLIYRRNPKT
GVPIGHWPVPESFWPDQNSPTLPPRTSHPVVKFSCTDCEPMVIDKLPFDKYELEPSPL
TOFILERKSPQTCWQVYSNSAKYSELCGHPFGYLKASTALNCVNLFVWPYNYVLLPL
LDDLFKVHRAKPILKWRQSFESVLKTMPPYYLGPLKKAVRMMGAPNLIADSMEYGLSY
SVISYLKKLSQOAKIESDRVIGSVGKKVVQEETGIKVRSSHGSLMSAYRKDFQOLLOGI
SEDVPHRLLDLNMKEYTGPQVALLNKDLKPQTFRNAYDIPRNLLDHLITRMRSNLLKS
TRRFLKGQDEDQVHSVPIAOMGNVQEYLKOVPSPLRELDPDQPRLHFFGNPFKLDKK
GMMIDEADEFVAGPQNKHKRPGEPNMQGTIPKRRRCMSPLLRGRQQNPVVNNHIGGKGP
PAPTTOAQPDLIKPLPLPLKISETTNDSIIHDVVENHVADOLSDDITPNAMDTEFSASS
PASLLERPINHMEALGHDHLIGTNDLTVGGFLENNEEPRDKEQCAEENIPASSLNKGKK
LMHCRSHEEVNTELKAQIMKEIRKPGRKYERIFTLLKHVQGSLQTRLIFLQNVIKEAS
RFKKRMLIEOLENFLDEIHRRANQINIHNSN"

BASE COUNT   1140 a   816 c   790 g   944 t
ORIGIN

Query Match   4.3%;   Score 86.4;   DB 40;   Length 3690;
Best Local Similarity   58.6%;   Pred. No. 3.5e-10;
Matches 150;   Conservative 0;   Mismatches 106;   Indels 0;   Gaps 0;

Qy 1581 aaaaaattaatgatgataaaatcaattaatgaaagaaagttcgaaagtttgggcaa 1640
        || |||||||||||||||||||||||||||||| | |||||||| | |||||||| ||
Db 2893 AAGAGGTCAATACTGAACTAAAAGCACAAATAATGAAAGAGATCCGAAAGCCAGGAAGAA 2952

Qy 1641 attatgaaagaattttcattttgcttgaagaaggtacaaggatctcatgaaagtcaagagac 1700
        | |||||||| |||| ||| ||||| |||||| || || |||||| |||||||| ||||||
Db 2953 AATATGAAAGAGAATCTTCACTTTACTGAAGCATGTGCAAGGCAGTTTACAACAAGACTAA 3012

Qy 1701 aatttgttgaatttacctcaaggaagcaggtttaaaaaaagttgtcttaattcgc 1760
        ||| || || |||| || || ||||||||||||||||||| ||| ||||||||||
Db 3013 TATTTTTACAAAATGTCATTAAAGAAGCATCAAGGTTTAAAAAAACGAATGCTAATAGAAC 3072

Qy 1761 aactcgagaaggcgcttaaagaaatagattcccactgccatctcaagaaaagttaagcaca 1820
        ||||| ||||| | ||| || |||||| |||| | | |||||||||||||||| |||||||
Db 3073 AACTGGAGAACTTCTTGGATGAAAATTCATCGAGAGCCAATCAGATCAACCATATTAATA 3132

Qy 1821 tgagaaaagatatt 1836
        ||| ||||| || |
Db 3133 GCAATTAAAAGAAAAT 3148

RESULT 6
LOCUS   AF141326   2110 bp   mRNA   PRI   17-JUN-1999
DEFINITION Homo sapiens RNA helicase HDB/DICE1 mRNA, partial cds.
ACCESSION AF141326
VERSION   AF141326.1   GI:5081649
KEYWORDS   .
SOURCE   human.
ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
           Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2110)
AUTHORS   Hoff,H.B. III, Oh,C. and Sell,C.
TITLE   Human DB1-1 homolog
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 2110)
AUTHORS   Hoff,H.B. III, Oh,C. and Sell,C.
TITLE   Direct Submission
JOURNAL   Submitted (07-APR-1999) Cell Cycle Control, Lankenau Medical
           Research Center, 100 Lancaster Ave, Wynnewood, PA 19096, USA
FEATURES   Location/Qualifiers
     source   1..2110
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /tissue_type="lung"
     CDS   <1..1849
                    /codon_start=2
                    /product="RNA helicase HDB/DICE1"
                    /protein_id="AAD39481.1"

RESULT 7

LOCUS   MMU57368   6346 bp   mRNA   ROD   30-MAY-1996
DEFINITION Mus musculus EGF repeat transmembrane protein mRNA, complete cds.
ACCESSION   U57368
VERSION   U57368.1 GI:1336627
KEYWORDS   .
SOURCE   house mouse.
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 6346)
  AUTHORS   Sell,C. and Hoff III,H.B.
  TITLE   Cloning of a novel mRNA regulated by the insulin like growth factor
            type 1 receptor
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 6346)
  AUTHORS   Sell,C.
  TITLE   Direct Submission
  JOURNAL   Submitted (01-MAY-1996) Christian B. Sell, Medical College of
            Pennsylvania, Gerontology, 2900 Queen Lane, Philadelphia, PA 19129,
            USA
FEATURES   Location/Qualifiers
     source   1..6346
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="DB1-1"
                     /clone_lib="TC-2"
                     /cell_line="R-"
                     /cell_type="fibroblast"
                     /tissue_type="whole embryo"

                     /dev_stage="fetal"
     5'UTR   1..693
     CDS   694..5757
                     /function="contact inhibition, anchorage dependence"
                     /note="notch like repeats; notch 2"
                     /codon_start=1
                     /evidence=experimental
                     /product="EGF repeat transmembrane protein"
                     /protein_id="AAB01338.1"
                     /db_xref="GI:1336628"
                     /translation="MSTALVAIRFLFTIVLKLFTISKOKOKPRLYEFISROHOKTEKE
                     GKQQLIIGFSARVELRNRRPGEVALLSRPPTPRRSLLSRMRPPAATCSPTFLRR
                     SPLSLDDYELVVSVCSCVWYYVNPASLPSYPVISNVEKLVPVKVICSAFMIFRCAI
                     TRSPCCSALAAPKMPALRPAGGGCCGLAVGGARPILQCCGGOEPCVNEGTCVTHNG
                     TGFCRCPEGFGLEEYCQHRDPCENRCQNGGTCVPQGMLGRATCRCAPGFTGEDCQYST
                     SHPCFVSRPCQNGGTCHMLSRDTYECTCOVGFTGKOCOWDACLSHPCGNGSTCTSVA
                     SOFSCKCPAGLIGKCEADINECDIPGRCOHGGTCLNLPGSYRCOCPQGFTGOHCDSP
                     YVPCAPSPCVNGGTCRQTGDTFECNLPGFEGSSTCERNIDDCPNHKCQNGGVCVDGV
                     NTYNCRCPPQWTGRVTVARVLGTWOTKLLARVGIVEYVTVGCLLARLECAVMLOFPGLE
                     WDSPVDSLPLKVCSOAPLGKKSSELRAGISTVLGWRAELSICQPVWHTOQVLCIVFE
                     REKNCTLIDTTNCAILKNPELTVYLLCQTAVLSQCCQOGWAGTLGDIIISTMSSFFVSFL
                     POWFKVSFFPTTDMSRSVVLKSLAVLKFLVELRFYVTHLFRLSDLYVGIRLDSAHVD
                     LFMTDFEISSKLVKKCKELNEIVFKLIMSKCWELKAVRHCVAVYILKCVLKCTIYDI
                     LNFTNRRDWVSEICIVDNKVFRIIEPDKLRLSLRTHIVGKNRDLPASCLDPTSTHSLT
                     FEYTHFQIKIKNNVLKMKREMNINSSRGIVGILFLRIGFVRLVRLFHQLPLAGSELTKEPFVGIRDY
                     TLGKKFLFLGKPAIIITITDGSKLTTSGVODELHLPNSPLAGSELTKEPFVGIRDY
                     LLLVLRLPGTMSVESEQLTGVPLDDSAITPMCEVTGGRSYSVCSPRMLNQCLESLVQK
                     VQSGVVINFEKRAGPDPPPAEAEGQODISRPEGSQPWHSCIKLIIYRMPYYLGPKVKH
                     PVPESFWPDONSPTLPPRTSHPVVKFPSCTDCEPMVIDKLPDKYLELEPSPLTQVSRR
                     SPQTCMQVYVSNSAKYNELGGHPFGVIKASTALTCVNLFVMPYNYPLPLLDDLFKVH
                     KAKPFLKWROSFESYLKTMPYYLGPLKAVRMMGAPNLIADSMEYGLSYSVISLKK
                     LSQQAKIESDRVIGSVGKKVVOETGIKVRSRSHGLSMAHRKGFQVLQGISEDVPHRL
                     DLNMKEYTGFQVALLNKDLKPQTFRNAYDIPRNLDHLTRMRSNLLKSTRKFLKGOD
                     EDQVHSVPIAQMGNQOEYLKQVPSLRELDPDQPRRLHTFGNPFKLDKDKKMMIDEADE
                     FVAGPONKHKRPGEPSMQGIPKRRCASPLLRGRRQSPAVNSHIGGKGPPAPMTOAQP
                     GLIKPLPHKEATNDSIVDDVVENHVADQLSSDMTPNAMDTEFLTSPPNLLEPSTNHT
                     EALGHEHLGNNDLTVGGFLENHEEPRNKEQSAEENIPASSLNKGKKLMHCRSHEEVNT
                     ELKAQIMKEIRKPGRKYERIFTLLKHVQGSLQTRLIFLQNVIKEASRFKKRMLIEQLE
                     NFLDEIHRRANQINIHNSN"
     3'UTR   5758..6346
                     /evidence=experimental

BASE COUNT   1837 a   1362 c   1369 g   1778 t
ORIGIN

Query Match   4.0%;   Score 80;   DB 12;   Length 6346;
Best Local Similarity   57.0%;   Pred. No. 1.1e-08;
Matches 146;   Conservative   0;   Mismatches 110;   Indels   0;   Gaps   0;

Qy   1581 aaaaaattaatgatgtataaaaatatcaattaatgaaagaagttcgaaagtttgggcaaa 1640
           |||||||||||| |||||||||||||||| |||| |||||||||||| |||||||||||
Db   5510 AAGAGGTCAACACTGAACTAAAAGCACAAATAATGAAAGAGATCAGAAAAGCCAGGAAGAA 5569

Qy   1641 attatgaaagaaattttcatttgcttgaaagagtacaaagatcctatgaaagtcaagagac 1700
           |||||||||||||| |||||||||||| |||||||||||||||||| |||||||||||||
Db   5570 AATATGAAAGAAATCTTCACCTTACTGAAGCATGTGCAAGGCAGTTTACAAACAAGGCTAA 5629

Qy   1701 aatttgttgaatttaccatcaaggaaggcagcaggtttaaaaaagttgtcttaattcagc 1760
           |||| |||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db   5630 TATTCTTACAAAATGTCATTAAGGAAGCCAATCAGTTTAAAGAAACGAATGCTAATAGAAC 5689

Qy   1761 aactcgagaaggcgcttaaagaaatagaattcccactgccatctcgaaaaagttaagcaca 1820
           |||| ||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db   5690 AACTGGAGAACTTCTTGGATGAAAAATTCATCGAAGAGACCAATCAAACACCATATTAATA 5749

Qy   1821 tgagaaaaagataatt 1836
           ||| |||||||||||
Db   5750 GCAATTAAAAGAAAAT 5765

RESULT 8

CVU89672

Query Match   4.28%;   Score 84.8;   DB 40;   Length 2110;
Best Local Similarity   58.28;   Pred. No. 8.2e-10;
Matches 149;   Conservative   0;   Mismatches 107;   Indels   0;   Gaps   0;

Qy   1581 aaaaaattaatgatgtataaaaatatcaattaatgaaagaagttcgaaagtttgggcaaa 1640
           |||||||||||| |||||||||||||||| |||| |||||||||||| |||||||||||
Db   1602 AAGAGGTCAATACTGAACTAAAAGCCCAAATAATGAAAGAGATCCGAAAAGCCAGGAAGAA 1661

Qy   1641 attatgaaagaaattttcatttgcttgaaagagtacaaagatctatgaaagtcaagagac 1700
           |||||||||||||| |||||||||||| |||||||||||||||||| |||||||||||||
Db   1662 AATATGAAAGAAATCTTCACTTTACTGAAGCATGTGCAAGGCAGTTTCAAACAAGGCTAA 1721

Qy   1701 aatttgttgaatttaccatcaaggaaggcagcaggtttaaaaaagttgtcttaattcagc 1760
           |||| |||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db   1722 TATTTTACAAAATGTCATTAAGGAAGCCAATCAGTTTAAAAGACGAATGCTGATAGAAC 1781

Qy   1761 aactcgagaaggcgcttaaagaaatagaattcccactgccatctcgaaaaagttaagcaca 1820
           |||| ||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db   1782 AACTGGAGAACTTCTTGGATGAAAATTCATCGAAGAGCCAATCAACCATATTAATA 1841

Qy   1821 tgagaaaaagataatt 1836
           ||| |||||||||||
Db   1842 GCAATTAAAAGAAAAT 1857

LOCUS   CVU89672   5726 bp   DNA   SYN   21-MAR-1997
DEFINITION Cloning vector pIREShyg, complete plasmid sequence.
ACCESSION   U89672
VERSION   U89672.1 GI:1899166
KEYWORDS   bicistronic.
SOURCE   unidentified cloning vector.
  ORGANISM unidentified cloning vector
            artificial sequence; vectors.
REFERENCE   1 (bases 1 to 5726)
  AUTHORS   Rees,S., Coote,J., Stables,J., Goodson,S., Harris,S. and Lee,M.G.
  TITLE   Bicistronic vector for the creation of stable mammalian cell lines
            that predisposes all antibiotic-resistant cells to express
            recombinant protein
  JOURNAL   BioTechniques 20 (1), 102-104 (1996)
   MEDLINE 96366236
REFERENCE   2 (bases 1 to 5726)
  AUTHORS   Lou,Y. and Holtz,A.
  TITLE   Direct Submission
  JOURNAL   Submitted (14-FEB-1997) CLONTECH Laboratories, Inc., 1020 East
            Meadow Circle, Palo Alto, CA 94303-4230, USA
COMMENT   This vector can be obtained from CLONTECH Laboratories, Inc., 1020
            East Meadow Circle, Palo Alto, CA 94303, USA. To place an order
            call (415) 424-8222 or (800) 662-2566, extension 1. International
            customers, please contact your local distributor. For technical
            information, call (415) 424-8222 or (800) 662-2566, extension 3.
            This sequence has been compiled from information in the sequence
            databases, published literature and other sources, together with
            partial sequences obtained by CLONTECH. If you suspect there is an
            error in this sequence, please contact CLONTECH's Technical Service
            Department at (415) 424-8222 or (800) 662-2566, extension 3 or
            E-mail TECH@CLONTECH.COM.
FEATURES   Location/Qualifiers
     source   1..5726
                     /organism="unidentified cloning vector"
                     /db_xref="taxon:45196"
     misc_feature   938..1233
                     /note="synthetic intron"
     intron   1273..1858
                     /note="internal ribosome entry site"
     CDS   1872..2906
                     /codon_start=1
                     /transl_table=1
                     /product="hygromycin B phosphotransferase"
                     /protein_id="AAB49979.1"
                     /db_xref="GI:1899167"
                     /translation="MDRSGKPELTATSVEKFLIEKFDSVSDLMQLSEGEESRAFSFDV
                     GGRGYKLRVNSCADGYKDRYVYRHFASAALPIPEVLDIGEFSESLTYCISRRAQGVT
                     LQDLPETELPAVLQPVAEAMDAIAAADLSQTSGFGPFGPQGIGQYTTWRDFICAIADP
                     HVYHWQTVMDDTVSASVAQALDELMLWAEDCPEVRHLVHADFGSNNVLTDNGRITAVI
                     DWSEAMFGDSQYEVANIFFWRPWLACMEQQTRYFERRHPELAGSPRLRAYMLRIGLDQ
                     LYQSLVDGNFDDAAWAQGRCDAIVRSGAGTVGRTQIRARSAAVWTDGCVEVLADSGNR
                     RPSTRPRAKE"
     gene   4730..5591
                     /gene="bla"
     CDS   complement(4730..5590)
                     /gene="bla"
                     /EC_number="3.5.2.6"
                     /function="ampicillin-resistance"
                     /codon_start=1
                     /transl_table=11
                     /product="beta-lactamase"
                     /protein_id="AAB49980.1"
                     /db_xref="GI:1899168"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNGGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTIIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW"
BASE COUNT   1328 a   1482 c   1518 g   1398 t
ORIGIN

RESULT   9
LOCUS   A49700   6802 bp   DNA   circular PAT   07-MAR-1997
DEFINITION Sequence 6 from Patent WO9610641.
ACCESSION   A49700
VERSION   A49700.1 GI:2303036
KEYWORDS
SOURCE   Influenza virus.
  ORGANISM Influenza virus
            Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
            unclassified Orthomyxoviridae.
REFERENCE   1 (bases 1 to 6802)
  AUTHORS   Hobom,G., Neumann,G. and Menke,A.
  TITLE   An attenuated vaccination virus, a method to make the virus and a
            pharmaceutical compositions comprising the virus
  JOURNAL   Patent: WO 9610641-A 6 11-APR-1996;
            BAYER AG (DE)
COMMENT   Other publication AU 3607695 960426.
FEATURES   Location/Qualifiers
     source   1..6802
                     /organism="Influenza virus"
                     /isolate="PHL1191"
                     /db_xref="taxon:11309"
BASE COUNT   1674 a   1695 c   1742 g   1691 t
ORIGIN

Query Match   3.4%; Score 69; DB 14; Length 5726;
Best Local Similarity   93.5%; Pred. No. 3.8e-06;
Matches   72; Conservative   0; Mismatches   5; Indels   0; Gaps   0;

Qy   1 ctcactatagggagacccacgctctggatccgagctcggatccactagtaacggccgccag 60
              |||||||||||||||||| ||||||||||||||||| ||||||||||||||||||||||||
Db   871 CTCACTATAGGGAGACCCCAAGCTTGGTACCGAGCTCGGATCCACTAGTAACGGCCGCCAG 930

Qy   61 tgtgctggaaagtgttc 77
              ||||||| |||| |||||
Db   931 TGTGCTGGAATTAATTC 947

RESULT 10
LOCUS   A58521   10443 bp   DNA   circular PAT   05-MAR-1998
DEFINITION Sequence 25 from Patent WO9638555.
ACCESSION   A58521
VERSION   A58521.1 GI:3714137
KEYWORDS
SOURCE   unidentified.
  ORGANISM unidentified
            unclassified.
REFERENCE   1 (bases 1 to 10443)
  AUTHORS   Bogaert,T., Stringham,E. and Vandekerckhove,J.
  TITLE   PROCESSES FOR THE IDENTIFICATION OF COMPOUNDS WHICH CONTROL CELL
            BEHAVIOUR, THE COMPOUNDS IDENTIFIED AND PHARMACEUTICAL COMPOSITIONS
            CONTAINING THEM AND THEIR USE IN THE CONTROL OF CELL BEHAVIOUR
  JOURNAL   Patent: WO 9638555-A 25 05-DEC-1996;
            BOGAERT THIERRY (BE)
COMMENT   Other publication AU 6123496 961218.
FEATURES   Location/Qualifiers
     source   1..10443

Query Match   3.4%; Score 68.4; DB 5; Length 6802;
Best Local Similarity   98.6%; Pred. No. 5.3e-06;
Matches   69; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   1 ctcactatagggagaccacgctggtaccgagctcggatccactagttaaacggccgccag 60
              ||||||||||||||||||||| ||||||||||||||||||||||| ||||||||||||||
Db   6697 CTCACTATAGGGAGACCCAAGCTTGGTACCGAGCTCGGATCCACTAGTAACGGCCGCCAG 6756

Qy   61 tgtgctggaa 70
              ||||||||||
Db   6757 TGTGCTGGAA 6766

BASE COUNT   2747 a   2731 c   2399 g   2566 t
ORIGIN
Query Match   3.4%;   Score 68.4;   DB 5;   Length 10443;
Best Local Similarity   98.6%;   Pred. No. 5.3e-06;
Matches   69; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   1 ctcactataggagagaccaccagcttggtacgcagctcggatcccactagtagtaacggccgccag 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 10349 CTCACTATAGGAGAGACCACCAGCTTGGTACGCAGCTCGGATCCCACTAGTAGTAACGGCCGCCAG 10408

Qy   61 tgtgctggaa 70
        ||||||||||
Db 10409 TGTGCTGGAA 10418

/organism="unidentified"
/db_xref="taxon:32644"

LOCUS   EVEI32038   3840 bp   mRNA   circular   SYN   28-JUL-1999
DEFINITION Expression vector pCDPI.
ACCESSION   AJ132038
VERSION   AJ132038.1 GI:5640088
KEYWORDS   AMP gene; beta lactamase; ColE1 origin of replication; multiple
            cloning site; SP6 promoter; SV40 origin of replication; T7
            promoter; xanthine-guanine phosphoribosyl transferase;
            Xanthine-guanine phosphoribosyl transferase gene.
SOURCE   Expression vector pCDPi.
  ORGANISM Expression vector pCDPi.
            artificial sequence; vectors.
REFERENCE   1 (bases 1 to 3840)
  AUTHORS   Zeng,B.J.
  TITLE   Mammalian Expression Vector for with fuse Xanthine-guanine
            phosphoribosyl transferase Tag
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 3840)
  AUTHORS   Zeng,B.J.
  TITLE   Direct Submission
  JOURNAL   Submitted (27-FEB-1999) Zeng B.J., Gene Engineering Center,
            Institute of Microbiology, Zhongguancun, Beijing, Beijing 100080,
            CHINA
FEATURES   Location/Qualifiers
     source   1..3840
                     /organism="Expression vector pCDPI"
                     /db_xref="taxon:90749"
     promoter   209..863
                     /note="CMV"
     promoter   864..882
                     /note="T7"
     misc_feature   882..984
                     /note="Multiple cloning site; HindIII, BamHI, BstXI,
                     EcoRI, NotI, XhoI"
     CDS   929..1387
                     /codon_start=1
                     /product="Xanthine-guanine phosphoribosyl transferase"
                     /protein_id="CAB51567.1"
                     /db_xref="GI:5640089"
                     /translation="MSEKYIVTWDMLQIHARKLASRLMPSEQWKGIIAVSRGGLVPGA
                     LLARELGIRHVDVICISSYDHDNQRELKVLKRAEGDGEGFIVIDDLVDTGGTAVAIRE
                     MYPKRAHFVTIFAKPAGRPLVDDYVVDIPQDTWIEQPWDMGVFVPPISGR"
     polyA_site   1649..1863
                     /note="BGH"
     promoter   2450..2775
                     /note="SP6"
     rep_origin   2644..2729
                     /note="SV40"
     gene   complement(2844..3704)
                     /gene="amp"
     CDS   complement(2844..3704)
                     /gene="amp"
                     /codon_start=1

Qy   61 tgtgctggaa 70
        ||||||||||
Db 10349 CTCACTATAGGAGAGACCACCAAGCTTGGTACCGAGCTCGGATCCACTAGTAACGGCCGCCAG 10409

Query Match   3.4%;   Score 68.4;   DB 14;   Length 3840;
Best Local Similarity   98.6%;   Pred. No. 5.3e-06;
Matches   69; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   1 ctcactataggagagaccaccagcttggtacgcagctcggatcccactagtagtaacggccgccag 60
        |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 796 CTCACTATAGGAGAGACCACCAAGCTTGGTACCGAGCTCGGATCCACTAGTAACGGCCGCCAG 855

Qy   61 tgtgctggaa 70
        ||||||||||
Db 856 TGTGCTGGAA 865

/product="beta-lactamase"
/protein_id="CAB51568.1"
/db_xref="GI:5640090"
/translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEQLGARVGY
IELDLNSGKILESRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLTTIGGPKELTAFLHNMGDHVTRL
DRWEPELNEAIPNDERDTIMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
EIGASLIKHW"
/note="ColE1"

LOCUS   PCDNA3ZEO   3986 bp   DNA   SYN   16-AUG-1995
DEFINITION Cloning vector pcDNA3ZEO DNA.
ACCESSION   X90639
VERSION   X90639.1 GI:949972
KEYWORDS   cloning vector; expression vector; multiple cloning site; plasmid.
SOURCE   synthetic construct.
  ORGANISM synthetic construct
            artificial sequence.
REFERENCE   1 (bases 1 to 3986)
  AUTHORS   Peters,H., Hundhausen,T., Kroenke.M. and Marget,M.
  TITLE   A new small sized high-level eukaryotic expression vector
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 3986)
  AUTHORS   Peters,H.
  TITLE   Direct Submission
  JOURNAL   Submitted (07-AUG-1995) H. Peters, Inst. f. Immunologie,
            Michaelisstr.5, D- 24105 Kiel, FRG
COMMENT   Related sequences: M21295 and K03104.
FEATURES   Location/Qualifiers
     source   1..3986
                     /organism="synthetic construct"
                     /plasmid="pcDNA3ZEO"
                     /db_xref="taxon:32630"
     misc_feature   1..2125
                     /note="cloning vector (pcDNA3) (Invitrogen)"
     misc_feature   889..994
                     /note="multiple cloning site (MCS)"
     misc_feature   2126..2796
                     /note="cloning vector (PZeoSV) (Invitrogen)"
     misc_feature   2797..3986
                     /note="cloning vector (pcDNA3)"

BASE COUNT   900 a   1057 c   1054 g   975 t
ORIGIN

Query Match   3.4%;   Score 68.4;   DB 14;   Length 3986;
Best Local Similarity   98.6%;   Pred. No. 5.3e-06;
Matches   69; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   1 ctcactataggagagaccaccagcttggtacgcagctcggatcccactagtagtaacggccgccag 60
        |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 871 CTCACTATAGGAGAGACCACCAAGCTTGGTACCGAGCTCGGATCCACTAGTAACGGCCGCCAG 930

RESULT 13

```
Qy      61 tgtgctggaa 70
            ||||||||||
Db     931 TGTGCTGGAA 940
```

```
Db     931 TGTGCTGGAA 940
```

RESULT 13

```
LOCUS       XXU43958      7290 bp    DNA      SYN      28-JAN-1999
DEFINITION  Cloning vector pRCMV-luc luciferase gene, complete cds.
ACCESSION   U43958
VERSION     U43958.1  GI:4097011
KEYWORDS
SOURCE      Cloning vector pRCMV-luc.
  ORGANISM  Cloning vector pRCMV-luc
            artificial sequence; vectors.
REFERENCE   1  (bases 1 to 7290)
  AUTHORS   Sakamoto,N., Ito,Y., Wu,G.Y. and Wu,C.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-DEC-1995) N. Sakamoto, Division of Gastroenterology,
            University of Connecticut Health Center, Farmington, CT 06030, USA
FEATURES             Location/Qualifiers
     source          1..7290
                     /organism="Cloning vector pRCMV-luc"
                     /db_xref="taxon:4541"
     mRNA            842..2870
                     /product="luciferase"
     CDS             993..2645
                     /codon_start=1
                     /transl_table=11
                     /product="luciferase"
                     /protein_id="AAD10138.1"
                     /db_xref="GI:4097012"
                     /translation="MEDAKNIKGPAPFYPLEDGTAGEQLHKAMKRYALVPGTIAFTD
                     AHIEVNITYAEYFEMSVRLAEAMKRYGLNTNHRIVVCSENSLQFFMPVLGALFIGVAV
                     APANDIYNERELLNSMNISQPTVVFVSKKGLQKILNVQKKLPIIQKIIIMDSKTDYQG
                     FQSMYTFVTSHLPPGFNEYDFVPESFDRDKTIALIMNSSGSTGLPKGVALPHRTACVR
                     FSHARDPIFGNQIIPDTAILSVVPFHHGFGMFTTLGYLICGFRVVLMYRFEEELFLRS
                     LQDYKIQSALLVPTLFSFFAKSTLIDKYDLSNLHEIASGGAPLSKEVGEAVAKRFHLP
                     GIRQGYGLTETTSAILITPEGDDKPGAVGKVVPFFEAKVVDLDTGKTLGVNQRGELCV
                     RGPMIMSGYVNNPEATNALIDKDGWLHSGDIAYWDEDEHFFIVDRLKSLIKYKGYQVA
                     PAELESILLQHPNIFDAGVAGLPDDDAGELPAAVVVLEHGKTMTEKEIVDYVASQVTT
                     AKKLRGGVFVDEVPKGLTGKLDARKIREILIKAKKGGKSKL"
BASE COUNT     1784 a   1800 c   1832 g   1874 t
ORIGIN
```

```
Qy       1 ctcactattaggggaaggaccacgcgcttggttaccggatccactagttaacggccgccag 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     872 CTCACTATTAGGGAGGAGACCCAAGCTTGGTACCGAGCTCGGATCCACTAGTTAACGGCCGCACAG 931
```

```
Qy      61 tgtgctggaa 70
           ||||||||||
Db     932 TGTGCTGGAA 941
```

RESULT 14

```
LOCUS       AR062871      4525 bp    DNA      PAT      29-SEP-1999
DEFINITION  Sequence 2 from patent US 5843770.
ACCESSION   AR062871
VERSION     AR062871.1  GI:5990562
KEYWORDS
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 4525)
  AUTHORS   Ill,C.R. and Gonzales,J.E.N.
  TITLE     Antisense constructs directed against viral post-transcriptional
            regulatory sequences
```

```
JOURNAL     Patent: US 5843770-A 2 01-DEC-1998;
FEATURES             Location/Qualifiers
     source          1..4525
                     /organism="unknown"
BASE COUNT     1115 a   1132 c   1168 g   1110 t
ORIGIN
```

```
Query Match             3.3%;  Score 66.8;  DB 5;  Length 4525;
Best Local Similarity  97.1%;  Pred. No. 1.2e-05;
Matches   68;  Conservative  0;  Mismatches  2;  Indels  0;  Gaps  0;
```

```
Qy       1 ctcactattaggggaaggaccacgcgcttggtaccgagctcggatccactagttaacggccgccag 60
           ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||| |||
Db    3070 CTCACTATTAGGGAGGAGACCCAAGCTTGGTACCGAGCTCGGATCCACTAGTTAACGGCCGCACAG 3129
```

```
Qy      61 tgtgctggaa 70
           ||||||||||
Db    3130 TGTGCTGGAA 3139
```

RESULT 15

```
LOCUS       I66494            7218 bp    DNA      PAT      28-DEC-1997
DEFINITION  Sequence 14 from patent US 5670367.
ACCESSION   I66494
VERSION     I66494.1  GI:2724471
KEYWORDS
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 7218)
  AUTHORS   Dorner,F., Scheiflinger,F. and Falkner,F.Gunter.
  TITLE     Recombinant fowlpox virus
  JOURNAL   Patent: US 5670367-A 14 23-SEP-1997;
FEATURES             Location/Qualifiers
     source          1..7218
                     /organism="unknown"
BASE COUNT     1944 a   1491 c   1486 g   1929 t   368 others
ORIGIN
```

```
Query Match             3.1%;  Score 62.2;  DB 5;  Length 7218;
Best Local Similarity   7.2%;  Pred. No. 0.00015;
Matches   31;  Conservative  225;  Mismatches 173;  Indels  0;  Gaps  0;
```

```
Qy    1588 taatgatgatataaaaattcaattaatgaaagaaagttcgaaagtttgggcaaaattatga 1647
           :::::   ::::: ::: :::::::: ::::::: ::: :: ::: :::::: :::::
Db    1494 TAGGCATCACTGTAATTACCTATCTATGCAAGTAGTTAAAGAGATAGAAGAATTTGGTAC 1435
```

```
Qy    1648 aagaatttcaatttgcttgaagaggtaccaaggatctatgaaggtcaagagacaaattgt 1707
           ::::: :::::::::::::::::::: ::::::::::::::::::::::::::::::::::::
Db    1434 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1375
```

```
Qy    1708 tgaattaccatcaaggaaggcaaggagttaaaaggttgttcttaattcaagcaaactcga 1767
           ::::: ::::::::: ::::::: :::::::::::::::::
Db    1434 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1375
```

```
Qy    1768 gaaggcgcttaaagaaaatagaatccccactgccatctccgaaaagttaagcacatgagaa 1827
           ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db    1314 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1255
```

```
Qy    1828 aagataattgtgttagtgcaaagaccaaggagaaacaagagacatatgctgtaggatgaa 1887
           ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db    1254 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1195
```

```
Qy    1888 caggttattgctgaagctccctataatcctgaaatgaagagaattcccttccagaagcta 1947
           :::::: :::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db    1194 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1135
```

```
Qy    1948 cgaaaaaggagctgttaaattaataaatctctgttagtaaagctgcaaaaaaa 2007
           :::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db    1134 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1075
```

Qy   2008 aaaaaaaaa 2016
      ::::::::::|
Db   1074 RRRRRRRA 1066

Search completed: March 24, 2000, 04:11:30
Job time: 8624 sec